Georg Boenn

Computational Models of Rhythm and Meter

VISIT…

Georg Boenn
Department of Music
University of Lethbridge
Lethbridge, Alberta
Canada

ISBN 978-3-319-76284-5 ISBN 978-3-319-76285-2 (eBook)
https://doi.org/10.1007/978-3-319-76285-2

Library of Congress Control Number: 2018942628

© Springer International Publishing AG, part of Springer Nature 2018

Preface

This book is the result of nearly thirteen years of my research in the area of rhythm, meter, and expressive timing. I hope it will be beneficial for composers, musicologists, musicians, software developers, as well as for the research communities working in these areas.

I wish to express my gratitude to the people who taught me, who shared their knowledge, and who gave advice, support and encouragement.

I would like to start with Clarence Barlow whose course *Musiquantenlehre* I took at the Musikhochschule in Cologne, Germany, in 1990. It was his intense one-week course, which set me onto the path of using computer algorithms for music composition and analysis ever since. The reader will find two of his formulas in this book, which are being applied to rhythm analysis and quantization.

Thanks also to John Fitch who supervised my Ph.D. thesis, and who first told me about the Farey Sequence; to Martin Brain as well, with whom I have had the pleasure to work with at the University of Bath, England.

Special thanks to Peter Giger and his wonderful book on Rhythm, from which I took inspiration for my shorthand notation. To my percussion colleagues at the University of Lethbridge, I say thanks to Adam Mason and Joe Porter for their support. Thanks to the Faculty of Fine Arts, and to the Dean, Ed Jurkowski, who gave me the space and the time for writing this book.

To my wife Daiva, I am ever grateful for her love, friendship, and beautiful companionship during all these years. I dedicate this book to our children.

I thank my editor, Helen Desmond from Springer Nature in London, UK, for her patience and advice. The plans for this book started some years ago in Bath, and despite of my move to Canada, she kept in touch.

Finally, this book would not have been possible without the many people of the open-source communities who create some of the most amazing music software tools. Especially, I would like to mention the Csound community, and the developers and users of Lilypond: Thank you so much.

Lethbridge, Alberta, Canada — Georg Boenn
April 2018

Contents

1 Introduction 1
References 6

2 Phenomenology of Rhythm and Meter 7
2.1 Causality 7
2.2 Definitions of Rhythm and Meter 8
2.3 Organic Form 10
2.3.1 The Cycle in Organic Form 11
2.3.2 Breathing 12
References 13

3 A Shorthand Notation for Musical Rhythm 15
3.1 Introduction 15
3.2 Overview of Rhythm Notation 16
3.3 Chunks of Musical Time: A Shorthand Notation for Rhythm 18
3.3.1 Rhythm and the Psychology of Chunking 18
3.3.2 Subdivisions 20
3.4 Examples 21
3.4.1 The Ewe Rhythm 23
3.4.2 Latin-American Music 24
3.4.3 Greek Verse Rhythms 25
3.4.4 Messiaen 25
3.4.5 Beethoven 27
3.4.6 Mussorgsky 28
3.4.7 Debussy 29
3.4.8 Polyrhythm 29
3.4.9 Conclusion of Examples 29
3.5 Conclusion 30
References 31

4 Partitions and Musical Sentences . . . 33
4.1 Introduction . . . 33
4.2 Integer Partitions . . . 34
4.2.1 Partitions into k Distinct Parts . . . 35
4.2.2 Partitions into Parts with an Arithmetic Progression . . . 36
4.3 Musical Sentences . . . 36
4.4 Asymmetric Sentences . . . 39
4.4.1 Stravinsky's Game with Metric Asymmetry . . . 39
4.4.2 Messiaen: The Birds as Teachers of Composition . . . 41
4.5 Measuring Metric Complexity . . . 44
4.6 The Resolution of Musical Sentences: Effects of Closure and Decline . . . 46
4.6.1 Shrinking Durations, or the Accelerando Technique . . . 47
4.6.2 Triangular Rhythmic Phrases using Primes . . . 48
4.7 The Sentence Algorithm in *Chunking* . . . 49
4.7.1 Seven Categories of Rhythmic Patterns . . . 52
4.7.2 Transcription of Patterns and the Complete Sentence . . . 53
4.8 Conclusion . . . 54
References . . . 55

5 The Use of the Burrows–Wheeler Transform for Analysis and Composition . . . 57
5.1 Introduction . . . 57
5.2 The BWT Algorithm . . . 58
5.2.1 The Inverse BWT Algorithm (iBWT) . . . 60
5.2.2 A Rhythm Analysis Program Using the BWT . . . 61
5.2.3 Fragmentation Modelling by Using the iBWT Matrix . . . 61
5.3 Conclusion . . . 62
References . . . 63

6 Christoffel Rhythms . . . 65
6.1 Introduction . . . 65
6.2 Christoffel Rhythms from Christoffel Words . . . 66
6.2.1 Operations on Christoffel Rhythms . . . 67
6.3 The Burrows-Wheeler Transform as a Tool for Rhythm Analysis . . . 69
6.4 Rhythms from Various Music Cultures . . . 71
6.4.1 Euclidean Rhythms . . . 77
6.5 Conclusion . . . 79
References . . . 80

7 The Farey Sequence as a Model for Musical Rhythm and Meter . . . 83
7.1 Introduction . . . 83
7.2 The Farey Sequence . . . 84

7.2.1 Building Consecutive Ratios Anywhere in Farey Sequences . . . 88
7.2.2 The Farey Sequence, Arnol'd Tongues and the Stern–Brocot Tree . . . 88
7.2.3 Farey Sequences and Musical Rhythms . . . 89
7.3 Filtered Farey Sequences . . . 92
7.3.1 Introduction . . . 92
7.3.2 Polyrhythms . . . 93
7.3.3 Rhythm Transformations . . . 100
7.3.4 Greek Verse Rhythms . . . 100
7.3.5 Filters Based on Sequences of Natural Integers . . . 103
7.3.6 Filters Based on the Prime Number Composition of an Integer . . . 104
7.3.7 Metrical Filters . . . 107
7.4 Conclusion . . . 110
References . . . 111

8 Models of Musical Meter, Temporal Perception and Onset Quantization . . . 113
8.1 Introduction . . . 113
8.2 Musical Meter . . . 114
8.2.1 Necklace Notation of Rhythm and Meter . . . 115
8.2.2 Meter and Entrainment . . . 117
8.3 Temporal Perception . . . 119
8.3.1 Shortest Timing Intervals . . . 120
8.3.2 The 100 ms Threshold . . . 121
8.3.3 Fastest Beats . . . 121
8.3.4 Slowest Beats . . . 122
8.3.5 The Perceptual Time Scale . . . 122
8.4 Onset Detection . . . 122
8.4.1 Manual Tapping . . . 122
8.4.2 Onset Data Extracted from Audio Signals . . . 124
8.4.3 Adjacent Interval Spectrum . . . 124
8.4.4 Is Knowledge of Onset Times Sufficient? . . . 126
8.5 Agogics . . . 127
8.6 Gestalt Theory . . . 128
8.7 Modelling of Neural Oscillations for Musical Meter . . . 131
8.8 Bayesian Techniques for Meter Detection . . . 132
8.9 Quantization . . . 134
8.9.1 Grid Quantization . . . 134
8.9.2 Context-Free Grammar . . . 135
8.9.3 Pattern-Based Quantization . . . 135
8.9.4 Models Using Bayesian Statistics . . . 136
8.9.5 IRCAM's KANT . . . 137

8.10 Tempo Tracking . . . 139
8.10.1 Multi-agent Systems . . . 139
8.10.2 Probabilistic Methods . . . 140
8.10.3 Pattern Matching . . . 140
8.11 Conclusion . . . 141
References . . . 143

9 Rhythm Quantization . . . 147
9.1 Introduction . . . 147
9.2 Grouping of Onsets into Duration Classes . . . 148
9.3 Quantization to a Metrical Grid . . . 151
9.4 Some Further Examples of Grouping . . . 155
9.5 Quantization of Onsets to a Filtered Farey Sequence . . . 156
9.6 The Transcription Algorithm . . . 157
9.6.1 Analysis Windows of Arbitrary Length . . . 157
9.7 Experimental Framework . . . 158
9.7.1 Test Material . . . 159
9.7.2 Distance Measurements . . . 159
9.8 Test Results . . . 160
9.8.1 Observations . . . 164
9.9 Conclusion . . . 171
9.9.1 Some Final Thoughts . . . 172
References . . . 172

Appendix A . . . 175

Index . . . 185

Chapter 1
Introduction

Computational Models of Rhythm and Meter presents several mathematical models and algorithms for the analysis and generation of musical rhythms and forms that range from the smallest patterns up to an entire musical piece. It is accompanied by an open source code project called *chunking*.[1] *Chunking* implements most of the algorithms that are presented in this book.

Music without rhythm is unthinkable. Music, like Theatre and Dance, needs time to perform. Rhythm is a universal principle. A large number of cycles and rhythmic patterns occur naturally within our own bodies and in the universe around us. Since all languages have their rhythms, music is often regarded as a language, a "rhythmic language of sound" (Giger 1993), although isolated musical sounds do not carry semantic meaning in the same way than words or sentences do in a language.

There is one universal law in music, and that is the law of impulse and release. A drum is hit with a stick, which means energy is brought into the system, and the drum head starts to vibrate in different modes. The acoustics of the drum will dictate how it will resonate afterwards until the sound fades away back into silence. This principle of impulse and release is one of arising energy, which might be even stable for some time, and subsequent dissipation of the same amount of energy that was brought in before. This principle can be found also on the scale of rhythmic patterns, musical tempo, musical sentences, up to entire musical pieces. A rhythmic pattern can create musical impulse and it can have a subsequent part, a resolution, which dissolves the previously created impulse. This triangular shape of beginning, middle, and end of musical energy over time is a reality, which we might also find represented in other visual metaphors, like the pyramid or a mountain.

We find ourselves constantly living inside overlapping cycles of cause and effect. The phases of arising, stable states, and decline are integral to our experience as living beings. The seasons of the year, day and night, the cycles of our daily routines, of sleeping, working, walking, and eating. The beating of our hearts, the continuous

[1] https://github.com/gboenn/chunking

thread of in-breath and out-breath, the blinking of an eye, the electromagnetic waves and impulses that go through our nerves and brains. In all of these rhythms and oscillations we find and ongoing transmission of causes and effects. It is, therefore, not really surprising that the same continuity of cause and effect is also present in music.

I will discuss this philosophy in more detail in Chaps. 2 and 4. I would like to keep these thoughts in mind when we will go over the other aspects of rhythm and expressive timing in musical performances, for example in Chaps. 8 and 9. Even if we are looking at rhythmic models from the abstract angle of mathematics in Chaps. 5, 6 and 7, one should not forget the perceptual aspects of impulse and release that are always coming into play when rhythms are being performed, and when they become an actual part of a musical experience. Therefore, I encourage the reader to try out for him- or herself the many rhythmic examples that are found in this book.

Computational Models of Rhythm and Meter deals with aspects of composition, with music analysis of rhythm and form, with transcription of music into notation, and with the analysis of recorded music performances, especially with the aspect of expressive timing. Parallel to writing this book, I developed a command-line tool called *chunking*, written in C++ and published under the GPL licence.[2] *Chunking* produces standard notation output using Lilypond,[3] and scores for the audio synthesis language Csound.[4]

A recurring concept in this book is that of interpreting rhythm patterns as words. A word is an ordered, finite or infinite sequence of symbols, or letters, taken from a non-empty alphabet, some finite set **A** (Glen 2012). For example, all words in this paragraph are finite sequences of letters taken from the finite English alphabet. An entire mathematical field called *Combinatorics on Words* emerged from investigations into the properties of words. I will demonstrate some applications of *Combinatorics on Words* within the realm of musical rhythm.

Outline of the Chapters

In Chap. 2, I discuss how the causes and effects of musical impulse and resolution influence a musical composition and how they guide our listening. This discussion is then linked to definitions of rhythm, polyrhythm and meter. The question about how rhythm is related to musical form will then lead us to discuss analogies between music and organic life, where we find similar structures of organization that are embedded within cycles of beginning, development and ending.

Chapter 3 presents a new shorthand notation for musical rhythms (SNMR). The SNMR is proposed as an easier way of notating rhythms, and for facilitating electronic storage in databases. It has certain benefits for music analysis because it reveals binary or ternary metric groupings in a compact form. In addition, it serves as a tool for creativity, for example in sketchbooks. In ASCII format, it can be sent to different backends for notation, database storage, and sound synthesis.

---

[2] https://github.com/gboenn/chunking

[3] www.lilypond.org

[4] www.csounds.com

Table 1.1 Symbols used for the shorthand notation for musical rhythm (SNMR)

ASCII symbol	Transcription	Binary Form	Category
.		1	unary
I		10	binary
:		11	
v		01	
–		100	
<		010	
w		001	
X		110	ternary
>		101	
+		011	
i		111	
H		1000	other
W		10000000	
~	e.g. H ~ I =	100000	tie
( )	e.g. (-) I =	00010	silence

The SNMR is a collection of ASCII symbols and it is based upon a small rhythmic unit called a pulse. The symbols refer to it in multiples of either two or three. Table 1.1 shows how the symbols can map to common practice notation. Note that the mapping can be changed depending on the musical context. For example, the mapping of the pulse symbol, the full stop, to an eighth note, , could change to use the 32nd note value instead: . The mappings of the whole set would then follow accordingly. Subdivisions can be realized by a combination of square brackets with a dividing factor.

Chapter 4 investigates asymmetric structures in the metric grouping of musical sentences. Integer partitions are being used as a tool in order to describe the relevant metric structures. The *sentence* algorithm in the software *chunking* models musical sentences as a hierarchy of phrases, patterns and rhythmic chunks. The design of the algorithms was influenced by recent studies in music phenomenology, and it makes references to psychology and cognition as well. I will also give a detailed analysis of the use of rhythm and meter in Igor Stravinsky's *Symphonies for Wind Instruments*. And, I will analyze the rhythms and metrical structures of two bird songs that have been recorded, transcribed and orchestrated by Olivier Messiaen for his orchestral work *Chronochromie*.

In Chap. 5, I investigate an algorithm for rhythm analysis that has been used so far mainly in bioinformatics and in file compression. The Burrows-Wheeler transform (BWT) works with a cyclic rotation of a pattern. The outcomes are listed underneath each other in lexicographic order. The last column of the list, read from top to bottom, is the Burrows-Wheeler transform of the input word. What makes the transform so interesting for musical rhythms is the inverse process of reconstructing the original input on the basis of the transformed word. When a musical phrase or sentence is repeated, one usually expects some sort of variation. The purpose of the variation is to generate a fresh impulse. A fragment generated via the inverse BWT of a rhythmic phrase can generate such an impulse. Another strategy that is covered here are the progressive shortening or expansion of rhythmic patterns. Inverse BWTs provide also a good tool for working with metric shifts of a musical rhythm.

Chapter 6 introduces a unique type of rhythmic patterns, which emerge from the application of a type of word that is known in the field of *combinatorics on words*, the so-called Christoffel word. Here we also use the Burrows-Wheeler transform in order to uncover substrings and rotations of the word. Words can be sequenced together via multiplication. With the help of these methods, numerous rhythmic patterns in different styles and from different cultural backgrounds can be described.

Chapter 7 gives a detailed account of the Farey Sequence. I am going to show that the Farey Sequence can serve as a general model for musical rhythm and meter. The Farey Sequence is a powerful tool to describe the structure of metrical subdivisions of beats, bars and even higher-level elements of musical form. The question is whether the Farey Sequence could explain the formal structure and rhythmic building blocks of entire sections of musical pieces, for which this chapter can give many evidences. Furthermore, the Farey Sequence is useful in order to analyze the rhythmic micro-structure of human performances of music. A Farey Sequence has the following definition, see Hardy and Wright (2008), p. 23:

A Farey Sequence F_n of order n is a list of fractions in their lowest terms between 0 and 1 and in ascending order. Their denominators do not exceed n. This means, a fraction a/b belongs to F_n, if

$$0 <= a <= b <= n \tag{1.1}$$

with $a, b, n \in \mathbb{N}$. In F_n, the numerator and denominator of each fraction is always coprime, and 0 and 1 are included in F_n as the fractions $0/1$ and $1/1$. For example:

$$F_5 = \left\{ \frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1} \right\} \tag{1.2}$$

A Farey Sequence can be used as an underlying grid structure for the analysis of musical performances. The elements of the Farey Sequence can serve as discrete time points to which the onset of sounds within a musical performance can be mapped. The Farey Sequence makes it then easy to translate the quantized time points to the specific values for musical notation. This process is known as quantization. In Chap. 9, I will

compare the outcomes of this process with the recorded musical performance, and I will compare it with a human transcription of that particular performance.

From the composer's point of view, the Farey Sequence is very interesting because one can model a diverse set of musical styles, for example Renaissance counterpoint and African Polyrhythm. Farey Sequences can translate directly into musical rhythms because their elements can determine the onset times of musical events on a variety of time scales. A wide range of rhythms can emerge from Farey Sequences. For example, for different values of n, F_n exhibits different rhythmic properties. It can encode musical polyrhythms or just any other conceivable compound rhythm that is aligned to a grid of small pulses, which are often a subdivision of a beat. To achieve this, various filters for Farey Sequences have been developed, which offer many possibilities for modelling rhythms and metrical structures.

Chapter 8 introduces the reader to important concepts of musical meter, and to the perceptual qualities of meter and expressive timing. How musicians perform music with expressive timing is being discussed from the angle of music history as well as viewed from the perspective of music psychology. I have included in this chapter the recent developments of beat and tempo tracking, as well as for onset detection, because these areas are dealing with challenges, which have to do with the human aspect of music making. Timing in a musical performance is never strictly mechanical. On the contrary, the possibilities of playing *rubato*, ahead of the beat, or more laid-back, are incredibly important aspects for the creation of a musical experience that reaches the listener in such a way that the experience is immediately understood as a fundamental *musical* effect.

Chapter 9 presents my algorithms for quantizing note onsets from recorded performances, and for the subsequent transcription of the rhythms into common practice notation. The quantization of note onsets prior to notation is necessary because human performances of music contain expressive timing. The tempo of a performance tends to have significant local fluctuations.

These tempo fluctuations are still a big challenge for beat tracking algorithms. Quantization is the mapping of performed note onset times and their durations to the onsets and durations defined by the score in common practice notation, or defined by learned metrical patterns. However, what if the score is unknown to the algorithm? What if the only information we have is a metrical tempo grid? Even with advances in sophisticated signal processing methods, the field still has to produce a complete system that would be able to compete with a human expert, especially when it comes to Classical music. See Benetos et al. (2013) for a good account of the state-of-the-art in Automatic Music Transcription.

What is notated in the score does not help much either. For example, tempo indications and rubato can only be indicated in words; nothing else in the score will explicitly show performers how to bend the musical time in an expressive way. What are the musical reasons for performers to slow down or to speed up? What are the musical circumstances that lead to changes in tempo? What musical factors lead to the decision to start playing in a certain tempo?

A score might only look like a collection of notes denoting their primary properties but what all performers strive to render is an aural image of the sounds. The aim

for them is to have a clear view of the connections between the sounds within the consciousness. A view of relationships is needed, which transcends the bare bones of musical facts that are described in a score. Clearly, where the listening experience is confronted with non-deterministic musical textures it will only be able to give statistical descriptions of the sounds. No transcription system is conceivable, in my view, that would be able to find the original score of a micro-polyphonic passage of Ligeti, for example.

My work was driven by the need for a reliable software tool that would be capable of rendering scores from musical performances in such a way that almost no editing of the score was needed. Specifically, the technology should find a truthful representation of the musical rhythms. One possible user scenario could be that a spontaneously created and improvised piece of music is recorded and then transcribed into a musical notation, which can be regarded as a truthful representation of this particular instance of a new piece that was created ad hoc, in realtime, freely improvised or perhaps being prepared from previous sketches. To capture rhythms truthfully is the challenging task here.

To achieve this goal, we had to develop and test various models of musical rhythms and meter. A new computational method of representing and analyzing such rhythms from various origins was needed as well. In Chaps. 8 and 9, the work is about a computational analysis of performed musical rhythms and to match the outcome with an appropriate form of rhythmic representation that could then be translated into a format for musical notation.

References

Benetos, E., Dixon, S., Giannoulis, D., Kirchhoff, H., & Klapuri, A. (2013). Automatic music transcription: Challenges and future directions. *Journal of Intelligent Information Systems*, *41*(3), 407–434.

Giger, P. (1993). *Die Kunst des Rhythmus : Professionelles Know How in Theorie und Praxis*. Mainz London New York: Schott. (ISBN 978-3-7957-1862-6).

Glen, A. (2012). *Combinatorics of Lyndon Words*, https://amyglen.files.wordpress.com/. Retrieved August 2015.

Hardy, G., & Wright, E. (2008). *An introduction to the theory of numbers* (6th ed.). Oxford: Oxford University Press.

Chapter 2
Phenomenology of Rhythm and Meter

This chapter is meant to give some philosophical thoughts on the subject of rhythm and musical form. For related discussions I would like to point the reader to Husserl (1980), Thakar (2011) and Boenn (2008). I will introduce in this chapter terminology that will recur throughout the book, and I am trying to give a greater perspective to these terms and why they are helpful for our discussion of rhythm and meter.

2.1 Causality

There is a causal continuity in music. One sound leads to the other, the next sound is the immediate consequence of what was sounding before. Each sound connects to the previous sounds in a network of cause and effect relations. A multitude of sounds brings forward the next sound. This sound, in turn, becomes the cause of future sounds, i.e. a single cause with many effects. As a result, a complex network of cause-and-effect relationships builds up in our minds over time. There is a continuous transmission of moments, a continuous transmission of impulse and release, which leads to a continuous stream of musical experiences. If the performance succeeds to balance out all forces exerted by the sounds over time, the ultimate goal of unity will be achieved and an experience of beauty will arise. This happens under the condition that the composition itself would create the right circumstances for an emergence of unity.

Rhythm is at the centre of the causal continuity that leads to a musical experience, because rhythm organizes the sequence of events. It determines when a new sound appears. On a larger scale, rhythm builds the musical form by arranging the sonic material in specific ways that speak to the psychology of the human mind. It plays a huge role in building the relationships between the sounds so that they have purpose and can become meaningful to the listener. Rhythm is to music what speaking and articulation is to a language. It is through rhythm that the other components,

especially melody and harmony, can come to life, so that they can generate their musical impulses, which lead to a musical experience. From the point of view of a composer, the problem seems to lie in the intuitive control of a multi-dimensional parameter-space. Each component in music, each force of impulse and release plays out over time. Therefore, all of music relies on some form of rhythm in order to materialize itself. From the point of view of the performers, who are also their own listeners, the problem is in the reduction of the movements through this multi-dimensional parameter-space into a perception of unity. This is a continuous act that plays out over the entire time of the performance. This act can be disturbed by forces that are external to the music, for example by the notorious mobile phone going off in the audience. Music is always structured in time through rhythm with the ultimate goal to unify the beginning, middle and ending of a piece, in such a way that all energies created by musical impulses have been fully released by equal amounts. If the release is too short, energy lingers on, and the experience remains in a state where one still expects for the sounds to lead to a closure. If the release is too long, then the sounds meander aimlessly because they are beyond the moment of closure already. Sounds that are configured to be in the phase of release become redundant after a particular point, which occurs when a full release had already been reached. The forces of impulse that were inserted from the very beginning reach eventually a point of maximum expansion, a climax point, after which opposite forces demand that the sounds will reach a point of full release in just the right amount of time. This balancing-act is part of the work of the performers and composers alike. There is a particular pattern to the experience of music: First the beginning, creation of impulse, eventually reaching a maximum expansion, the climax, then release and ending. I call this pattern a *musical triangle*.

2.2 Definitions of Rhythm and Meter

A rhythm emerges from a measured succession of sounds in time. It happens on the basis of a quasi iso-chronic pulsation. I say *quasi isochronic* because, in the reality of a performance, the tempo of pulsation might vary quite significantly. This is because the tempo of the pulsation is an important musical factor, which helps the mind to make sense of the rhythm. The tempo of the pulse is a condition for the perception of the unfolding impulse and release, including the perception of all phenomena that transmit these forces, at the right moments in time. In order to make sense of a series of sounds, a listener needs to focus the mind on an extended period of time, which is called the present moment. And, sounds can combine into music only when the mind becomes aware of them, i.e. when it is not distracted by other objects, or by thinking. There are some perceptual constraints that influence the choice of tempo made by music performers (London 2012). We will discuss them in Chap. 8 of this book. The choice of tempo should help the listeners to stay in the present moment and to become aware of the way how the sounds connect with each other. If the tempo is too fast, listeners may not grasp all of the musical phenomena; if the tempo is too slow, then the sounds cannot connect within the present moment and the unifying musical experience cannot establish itself.

A meter is a cyclic reiteration of a simple rhythm. Meters can have isochronous or non-isochronous beats.[1] A rhythm can become a polyrhythm. A polyrhythm is the superposition of rhythms that emerge from patterns of duration, accentuation, grouping, phrasing, tone colour, pitch, or change of harmony. The alignment of these patterns with the downbeat of a specific meter firmly establishes the perception of a meter.

A rhythm played in a certain meter becomes a polyrhythm via superposition and combination with the meter, which is itself regarded as a rhythm. A change of the meter, whilst keeping the same rhythm in combination, changes the polyrhythm. Shifting the metric position of that rhythm relative to the same meter changes the polyrhythm. We will see examples of metric shift in Chaps. 3 and 5.

Even without a meter, for example in Sub-Saharan music of West Africa, if a pattern of sound durations is combined with an accentuation pattern, a polyrhythm is created. If the accentuations change their positions, then the polyrhythm changes as well.

Musical impulse is the insertion of energy into a single sound or into a whole sequence of sounds. This can happen in various ways, by using dynamics, pitch, intervals, harmony, tone colour and, of course, through rhythm, i.e. through the relationships between sound durations. If the number of short durations in succession increases then it leads to more musical impulse. If the number of longer durations surpasses the shorter ones then it leads to a resolution or decline of musical impulse.

Syncopation occurs when a rhythm with a relatively high musical impulse is combined with a meter with relatively low musical impulse. A rhythmic impact value is high when the ratios of the durations of a rhythm involve higher prime factors in their denominators. A rhythmic impact value is relatively low when the rhythm is in close alignment with the regular beats of the meter. Then, the duration ratios have denominators with relatively small prime factors.

An important aspect of rhythm perception is the ability of the mind to correlate a polyphony of sound events. A random sequence of events can be perceived as they are, like anything else in present moment awareness, but a certain condition has to be met in order for them to be correlated as a musical rhythm. This condition is the alignment of the sound onsets to an underlying grid of smaller pulsations. These pulsations do not have to be manifest as real sounds, but they can be present in the mind of the listener. Correlation means the connection of all sound events to a common frame of reference. In the case of rhythm, this frame of reference is an underlying iso-chronic pulsation that allows for relationships to be built. These relations are the rhythmic durations combined with the musical forces of impact and resolution that act between them.

The durations of sounds in music have a measurable, physical amount of time. The mind interacts with the physical durations and with all other properties of sound simultaneously, and it is capable of forming relations that bind together the sounds into a unified experience.

---

[1] Meter will be covered in-depth in Chap. 8, Sect. 8.2.

Randomly timed sound events exclude themselves from correlation if they fall too often in-between the onsets of a mentally predictable grid of pulsations. If no pulsation is present, then the awareness will focus on the other parameters of sound in order to grasp an understanding of the succession of sounds in time.

2.3 Organic Form

In what sense can a piece of music have an *organic* form? How would we define organic form? "Used in the aesthetic sense, form means that a piece is *organized*; i.e. that it consists of elements functioning like those of a living *organism*." Arnold Schoenberg emphasized the words organizing and organism in his definition of musical form (Schoenberg 1985). In this view, music can show similarities to a natural organism, a living being. Sounds are being organized into a piece of music. The word organism goes back to the Greek words 'organismos' and 'organon', whose meanings are 'instrument', 'tool' and 'sense of apprehension'. These meanings underline the sense of purpose that we give to the sounds within a piece of music. According to biology, the characteristics of an organism, a living being, are: It consists of parts joining and working together, with one or more cells and organs (organelles) that complement each other. Small parts build up larger, more complex parts, which form a functional system with a hierarchy and complex interactions. An organism is capable of reaction to stimuli, and reproduces itself. It grows and develops into its maximum extension. It consistently tries to live in homeostasis, which is the process of self-regulation in order to achieve a state of maximum stability and balance. Upon death, the organism decomposes and its constituent elements fall apart and return back into life-less matter that can be consumed, perhaps, by other organisms.

The shape and the purpose of musical forms and their constituent elements are: Introduction, development, and closure. In other words, the unfolding of musical form follows the organic scheme of arising, growth to maximum degree, followed by decline. And, the other meaning of organic form is that music is perceived through the senses and specifically talks to our senses. The sense faculties make use of organs; the mind perceives the world and evaluates experiences that reach the mind through the organized sense faculties.

We are multicellular organisms with trillions of cells that specialize in functionality and purpose. Throughout human history, we seem to be on a path to a more and more detailed understanding of nature. We also seem to have a certain preference to produce things in life, and in art, that are behaving in a similar way to how an organism works. We create objects that might have a structure that is similar to living beings, with a defined purpose of all parts, with a hierarchy of parts; objects that are subject to change (reaction to stimuli), which can happen either in a positive or in a negative direction, i.e. development and growth versus decline, decay and destruction. This constant change shows us that everything within us, and outside of us, is a process. It shows that nothing is fixed, and that the homeostatic states of balance can only be temporary (Langer 1953).

Another analogy is, of course, that musical works come into being during a performance in the medium of time; they begin, last, and end, just like an organism

comes into being through birth, lives through its life-time, and then dies. Therefore, in discussions about organicism in music, for example (Watkins 2017), one should acknowledge that all beings have to die. Decay and death are necessary parts of how the world works. The point here, of course, is that the aspect of decline, closure and death also applies to sound phenomena, and to musical form as well.[2] The idea that there is some unbound, or 'eternal' growth in the world, and that music is somehow capable of reflecting this, cannot be true. This is not how life, and music, in reality works. The analogy between organic life and music holds particularly strong when it comes to the matter of death. Sounds and musical experiences always come to an end at some point. One says that a string or an echo dies out. Instead of neglecting the finite nature of all that is inside and around us, one should acknowledge that everything in this world is subject to change.[3] There is no such thing as an infinite sound on this planet. Beginning, development and ending are inherent in the physicality of sound. They are also inherent in the act of music making and all discussions of musical form are usually concerned about it (Caplin et al. 2010). The Romanian conductor Sergiu Celibidache once said that "after a symphony is finished, all the notes are gone, what remains? The relationships." (Fischer et al. 1986) This statement also points to the fact that every time we are making a musical experience in the real world, there is an active part in the listener who consciously or subconsciously always seeks out for connections between the experiences. After a while, the mind becomes entrained by recurring patterns. Relationships are reinforced, patterns are recognized, and out of them categories are being formed.

2.3.1 The Cycle in Organic Form

Another characteristic pattern of the process of life is the cycle. Giving and taking, exchanges of matter, is organized in a cyclic fashion. Exchange of matter can happen on the level of organs inside a body, or on the level of entire beings, in that they might join, live or work together, and complement each other. We are breathing in oxygen generated by green trees and plants, breathing out carbon-dioxide for the trees and plants to breath in. This constant exchange of matter is a reality. In the interaction between organisms that produce oxygen and those organisms that are in need of oxygen, there is a functional organic connection between the two groups of organisms.

Schoenberg (1985): "A well-balanced melody progresses in waves, i.e. each elevation is countered by a depression. It approaches a high point or climax through a

[2]Schoenberg (1985): "Rhythm is particularly important in moulding the phrase. It contributes to interest and variety; it establishes character; and it is often the determining factor in establishing the unity of the phrase. The end of the phrase is usually differentiated rhythmically to provide punctuation."

[3]Langer (1953, p. 99): "Because we are organisms, all our actions develop in organic fashion, and our feelings as well as our physical acts have an essentially metabolic pattern. Systole, diastole; making, unmaking; crescendo, diminuendo. Sustaining, sometimes, but never for indefinite lengths; life, death."

series of intermediate lesser high points, interrupted by recessions. Upward movements are balanced by downward movements; large intervals are compensated for by conjunct movement in the opposite direction. A good melody generally remains within a reasonable compass, not straying too far from a central range."

Certain visual metaphors, so-called Image schemas, have been proposed for music analysis by Brower (2000), who refers to the works of Johnson (1987), and Lakoff (1987). Growth, balance and decline are features of all organic life. Therefore, in addition to the above, I developed the image schema of a musical triangle that represents growth, balance and decline, the features of all organic life, which is cyclic in nature. Musical sentences are structures, where smaller parts with defined purposes join together to form a larger structure. The triangle shape can be understood as the rise and fall of musical energy over time via the continued formation of impulse and subsequent resolution (Thakar 2011; Brower 2000). Or, the triangle shape can also have the meaning: The lengths of smaller units and phrases who are gradually increasing and decreasing over the total length of a sentence. Musical triangles over a number components have an order that is governed by the principles growth (ascending) and decline (descending). The triangle is formed by combining growth and decline into one larger unit over time. In his works on music analysis, Olivier Messiaen, for example in Messiaen (2000), used the terms of anacrouse, accent and désinence (anacrusis, accent and ending). When the three terms can be applied to a musical phrase or sentence, we can find a musical triangle to structure this sentence, and to split it up into its components, see also Chap. 4.

2.3.2 Breathing

If the body is at rest, the ratio of the length of an in-breath over the length of an out-breath is approximately 1:2. The total length of one breath can easily be between 8 and 10 s. If we equate one half of the breath cycle to the psychological extend of the present moment in time, then the psychological present can be somewhere between 3 and 7 s long, with the in-breath representing the shorter leg of the proportion and the out-breath representing the long leg. One moment is the in-breath, the next moment is the out-breath. The breath is the thread on which our entire life is hanging. In music we find a real application and, at the same time, also a symbolic reflection of the breath. It has a real impact, because singers and wind instrument players have to use the breath in order to produce sounds. It can be also a symbol in the sense that longer time structures follow the existential core of breathing in, and breathing out. In the last chapter of this book, I will present test results of my algorithms for rhythm quantization and transcription. I am using two of Glenn Gould's recorded performances as test material. Gould is of course famous for his singing lines in synchrony with the sounds of his piano playing. For a detailed account on this aspect of Gould's performance, see Sanden (2013), pp. 44–64. In these tone documents we might remind ourselves about how important the breath is when making music, and especially on an instrument like the piano, whose mechanics are designed to serve as extensions of the human body.

References

Boenn, G. (2008). The importance of Husserl's phenomenology of internal time-consciousness for music analysis and composition. In *Proceedings of the ICMC 2008* (p. 4). Belfast: ICMA, Ann Arbor, Michigan (ISBN 0-9713192-6-X). http://hdl.handle.net/2027/spo.bbp2372.2008.157.

Brower, C. (2000). A cognitive theory of musical meaning. *Journal of Music Theory*, *44*(2), 323–379.

Caplin, W. E., Hepokoski, J., & Webster, J. (2010). *Musical form, forms and formenlehre: Three methodological reflections*. Leuven: Leuven University Press. (ISBN 978-9058678225).

Fischer, M., Holland, D., & Rzehulka, B. (1986). Gehörgänge. Zur Ästhetik der musikalischen Aufführung und ihrer technischen Reproduktion. In (chap. Musik verschwindet. Gespräch der Autoren mit Sergiu Celibidache 1985). Peter Kirchheim. (ISBN 3874100162).

Husserl, E. (1980). *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins*. Tübingen: Niemeyer.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press. (ISBN 978-0226403175).

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press. (ISBN 978-0226468044).

Langer, S. K. (1953). *Feeling and form: A theory of art developed from philosophy in a new key*. London: Routledge and Kegan Paul.

London, J. (2012). *Hearing in time. psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Messiaen, O. (2000). *Traité de Rhythme, de Couleur et d'Ornithologie* (vol. 6). Alphonse Leduc. (ISBN 978-2856890561).

Sanden, P. (2013). Liveness in Modern Music: Musicians, Technology, and the Perception of Performance. New York and London: Routledge. (ISBN 978-0415895408).

Schoenberg, A. (1985). *Fundamentals of musical composition*. London: Faber and Faber. (ISBN 0571092764).

Thakar, M. (2011). *Looking for the harp quartet: An investigation into musical beauty*. Rochester, NY: University of Rochester Press. (ISBN 978-1-58046-346-1).

Watkins, H. (2017). Toward a post-humanist organicism. *Nineteenth-Century Music Review*, *14*, 93–114.

Chapter 3
A Shorthand Notation for Musical Rhythm

3.1 Introduction

This chapter presents a complete set of symbols to be used as shorthand notation for musical rhythms (SNMR). I will use this shorthand notation of rhythmic patterns to illustrate nearly all of the musical examples in this book. Its compact form not only allows me to quote rhythms within the running text, it also helps to visualize metric groups of various sizes in a clear and concise way. Based on the work of Giger (1993), I have extended his small set of symbols into a comprehensive system for notation. The notation's use of cognitive chunking enables musicians and composers to notate rhythms faster than usual, for example, when compared with common Western score notation. The only constraint of the proposed system is that it is dependant on a common underlying pulsation. Whilst our common notation system creates binary subdivisions that are based on the whole note, the proposed shorthand notation is built upon a sign for a small pulsation, and has signs for ternary as well as binary groupings that are multiples of a common small value. The system also frees the user from the use of rigid bar and meter structures, i.e. it does not imply regular changes of strong and weak beats. The method has value for music analysis because it leads to a transparent grouping of basic rhythmic elements, and it reveals their metric meaning within the surrounding context. The system can easily express variations of a rhythmic group within the same meter. The structuring of rhythms in groups is very useful for the analysis of musical impulse and release, i.e. tension and relaxation, which, as we have learned in the previous chapter, are an essential feature of the musical experience.

Composers and musicians need a system to notate music quickly by hand, but common practice notation can be quite complex. Therefore, it can slow down the speed of handwriting, which, in the case of composition or dictation, is then unable to keep up with the pace of musical events. The proposed system offers a solution to that problem, but it is restricted to the notation of rhythms. It has not been developed with the aim to encode pitch information, however there is a possibility to combine shorthand rhythm notation with the encoding of pitch in common standard notation.

A shorthand notation for rhythm is useful as an intermediate stage of composition, before a score is written down in common practice notation. It is also useful at the rhythm input stage of music typesetting programs. Our software, *chunking*, uses shorthand notation as its input format. In addition, it has several algorithms, which produce shorthand notation in ASCII format. *Chunking* can translate SNMR into a *lilypond*[1] script, which then can be used to render a score. For musicians, the shorthand can raise the awareness for rhythmic grouping, articulation and form, whilst becoming independent of meter and bar lines. For the musicologist it is a useful tool for music analysis, to reveal commonalities and differences between rhythmic patterns. The simple format makes it also ideal for database searches. The current distribution of *chunking* contains an *sqlite* database[2] of rhythm patterns. Within a large catalogue of rhythms, the user can, for example, search for common substrings at the beginning, middle, or end of patterns.

In the first section, I will give a brief overview of various approaches to rhythm notation. Then I will present a coherent system of symbols that extends and modifies Giger's *rhythmoglyphs* (Giger 1993, p. 174). Several well-known music examples written in this system shall demonstrate its general purpose, speed and versatility. The chapter concludes with a summary of the results.

3.2 Overview of Rhythm Notation

Many approaches exist to the notation of rhythm and meter. The reader is referred to Sethares (2007), London (2012), Toussaint (2004), and Arom (1991) for detailed discussions. Current Western notation is based on the concept of meter, which is a regular grid formed on top of a constant underlying pulsation. Musical meter implies a ranking of the individual pulses within a bar. The importance and statistical weight of these pulses is directly dependant upon their location (Barlow 2008). For example, the first pulse of a bar in any meter usually has the strongest accent and it also carries the highest probability for an event to occur at that point. In order to place notes onto a metric grid, Western rhythm notation uses a set of symbols to denote integer multiples or fractions of a standard note length. It uses a relatively small set of integer ratios, as well as ties, rests, and articulation signs in order to vary note durations.

Musical meters, for example 4/4, 3/4, 5/4, or 12/8, repeat in cycles. The repetitive character of meter has been known also in other musical cultures for a long time. For example, already in the 13th century, Arab music theory developed a so-called necklace notation (Demaine et al. 2009; Sethares 2007; Chew and Rastall 2001). Rhythmic patterns and metric structures are represented by black and white circles on a ring. A black circle means a note onset and white circles count the space between notes. The total number of circles on the ring represents n pulses of a complete metric cycle. This type of notation has been revisited in recent publications on rhythm and meter (Sethares 2007; London 2012). Closely related to necklace notation is the

[1] http://lilypond.org.

[2] https://www.sqlite.org.

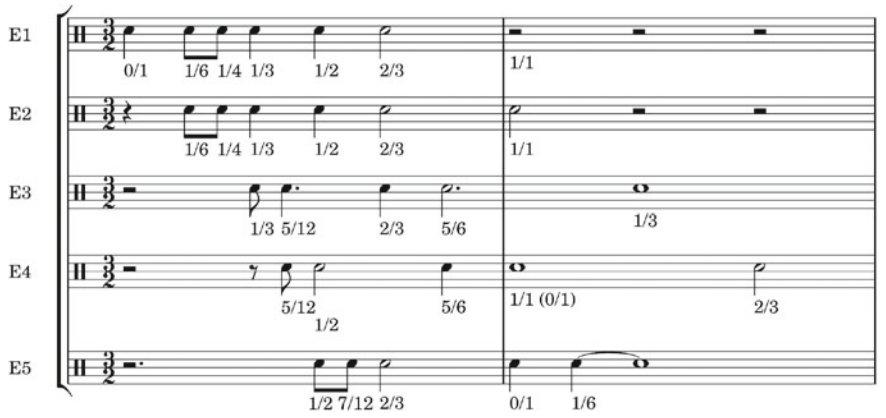



Fig. 3.1 A polyphonic score where the note onsets within a bar coincide with the *filtered* Farey Sequence $F_{12} = \{\frac{0}{1}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{5}{12}, \frac{1}{2}, \frac{7}{12}, \frac{2}{3}, \frac{5}{6}\}$

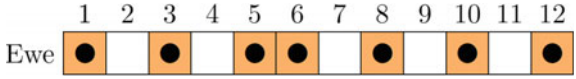


Fig. 3.2 Example of box notation: The 12-beat Ewe rhythm

Table 3.1 The set of Hindu symbols reported by Messiaen for the notation of the Deci-talas

Name	Symbol	Transcription
laghu	I	
guru	S	
pluta	S ‿	
druta	O	
laghu plus virama	I^c	
druta plus virama	O^c	

box notation, see Fig. 3.2. From this representation, binary and integer sequences are easily derived as well. A binary sequence for a rhythm uses '1' to denote a note onset and '0' to denote the time between onsets. A sequence of integers is yet another form of representation, and it counts the number of equidistant pulses between note onsets including the first note onset. This practice is also used in India (Giger 1993, p. 117). The specific Hindu notation reported by Messiaen (1995) also uses a small set of symbols, see Table 3.1. Furthermore, in Chap. 7, I propose the use of filtered Farey Sequences to represent rhythm and meter. They are particularly useful in order to apply weighted probabilities to the different onset locations within a metrical grid, see Fig. 3.1.

The main issues for a shorthand notation system are general usability, speed and flexibility. When it comes to selecting a shorthand notation for any form of rhythm

that is based on a common pulsation, all of the above methods are problematic. One of the main difficulties is the amount of counting that is involved in box, binary, integer or necklace notation. A unit or slot in these forms of notation usually represents a subdivision of a beat. Therefore, in order to notate, the user has to focus the attention on a low level of the metric structure and the tempo of the subdivision can be very fast. For example, a sixteenth note within a tempo of $quarter = 120$ BPM has a tempo of 480 BPM.[3] Also, a change of the subdivision of the beat, for example from eighth notes to eighth note triplets, is very difficult to represent due to the fixed length of the beat subdivision. The Hindu notation in Table 3.1 represents a very specific style of rhythm. In many of his compositions Messiaen has used some of the 120 deci-talas that were written in this form in Śārṅgadeva's Saṅgita Ratnākara (Bruhn 2007). Note that the symbols represent single note values, and because of this limitation, they are not general enough for our purpose.

Filtered Farey Sequences are a mathematical model for the representation of rhythm and meter. In Chap. 7, we will see that all rhythms, which are based on an underlying pulsation, can be represented by a filtered Farey Sequence. This includes structures that contain nested tuplets, i.e. integer subdivisions of a note value. But, they require the computation of ratios of note durations, and therefore, Farey Sequences have their role as a tool for music analysis, but they cannot function as a form of shorthand notation. Traditional Western notation is very precise, but often requires a combination of different symbols to encode a single note duration. It can be therefore quite slow for composing, sketches or for real-time dictation.

The problem is the demand for speed when capturing musical rhythms, handwritten or typed, combined with the demand for precision, and with the requirement to be easy to read. The proposed solution to the problem employs a technique called *cognitive chunking*, meaning that one can represent often recurring rhythms by one special symbol. Cognitive chunking is a well known psychological method to increase the number of elements that can be held in short-term memory (Gobet and Simon 1996; Miller 1956). The key concept is to find a memorable symbol that represents a rhythm that is either two, three or four pulses long. The set of symbols then forms a code that can *compress* the amount of rhythmic information more efficiently than in all of the above cases. This code can significantly speed up the notation process.

3.3 Chunks of Musical Time: A Shorthand Notation for Rhythm

3.3.1 Rhythm and the Psychology of Chunking

The Swiss percussionist Peter Giger introduced a handful of symbols for the notation of simple rhythmic patterns, which he calls *rhythmoglyphs*,[4] see Table 3.2. These symbols are easy to understand and to execute regardless of the level of knowledge

[3] beats per minute.

[4] Glyph means letter or symbol. A *rhythmoglyph* is a letter or a symbol representing a rhythm.

Table 3.2 Giger's rhythmoglyphs are representing five basic chunks of rhythm. Within the binary form, '1' denotes an onset, whereas '0' denotes silence

Symbol	Transcription	Binary Form
I		10
X		110
△		1110
□		11110
○		111110

Table 3.3 Complete set of six permutations of a pattern formed by three elementary chunks. The pattern has a cycle length of $n = 9$. The pattern is based on a template, which is the partition of 9 into $2 + 3 + 4$. Each slot of this template is then filled by the chunk of the corresponding length. There is only one choice per slot possible, because this is coming as a restriction from within Giger's set of symbols, see Table 3.2

Pattern	Transcription	Binary Form
I X △		101101110
I △ X		101110110
X I △		110101110
X △ I		110111010
△ X I		111011010
△ I X		111010110

of Western common notation. They are intuitive to use because the number of lines for drawing a symbol is equal to the number of sounding beats (except for the circle). With these elementary patterns, which I call *chunks*, one is able to construct a systematic table of combinations that invite the user to improvise and to invent new combinations by permutation, see Table 3.3. The resulting phrases are in alignment with a steady pulsation in time, and they do not necessarily invoke a sense of regular meter in the meaning of Western music traditions (London 2012; Arom 1991). The combinatorial power of Giger's approach is impressive. It makes it easy to create shorter and longer rhythmic phrases without thinking in terms of bars and meter. Instead, to play with these chunks seems closer to the way the mind works (Gobet et al. 2001; Miller 1956). It leads to a rich and embodied rhythmic experience. This method offers a huge range of possibilities for rhythmic variation. For example, after choosing one of the combinations of chunks, the performer is invited to swap their positions, which instantly creates new patterns. Some examples are given in Table 3.3. Furthermore, through repetition and sequencing of rhythmoglyphs one is able to build up higher level structures of climax and release, very similar to the process of thematic variation.

Within the context of music notation, Giger's ideas free the notation from bar lines and metric templates and it shifts the attention towards rhythmic groupings that can be re-assembled together in many playful ways. Through the concept of cognitive chunks and by practising the execution of those simple rhythmic patterns it is also relatively easy to leave behind the idea of counting beats, such as '1-and-a-2-and-a-3-and-a' in a 9/8 meter, for example. This freedom in turn allows the player to concentrate on the movements, on accentuation and on the 'groove' of the phrase. Continued practising and growing expertise will lead to bigger chunk sizes that can be held in working memory (Gobet and Simon 1996). That means that the length of rhythmic phrases that can be memorized and recalled will grow, and therefore it is likely that musical abilities and improvisational skills will increase as a consequence. Although this system already leads to a great number of different patterns, it is not covering every possibility for a rhythmic pattern. For that reason, I extended Giger's set of chunks with some more symbols, see Table 3.4, in order to have a complete set of unary, binary and ternary chunks available. In addition, following Giger's idea, silence is expressed by putting round brackets around the symbol. With this new set it is very easy to notate *any* kind of complex rhythmic phrase. Note the use of different signs for △, □, because they are not part of the printable ASCII codes 32-127, nor are they included in the extended set of ASCII codes 128-255. Therefore, in order to enhance the practicality of the shorthand notation on the computer - bearing in mind that it should be easy for the symbols to be typed - I am using `:I` and `iI` respectively. Various subdivisions of a unit, also known as tuplets, can be expressed via square brackets surrounding the symbols, for example `I[3III] I[3III] II I`. The divisor of the symbols inside the brackets is given as a rational number on the left-hand side. The following Table 3.5 gives an overview of the most common subdivisions and their transcription into shorthand notation.

The pulse lengths 2 and 3 are represented as `I` and `-`. The latter is the sign for 2 rotated by 90°. The orthogonal position helps to distinguish binary from ternary chunks. The rhythmic constellation of 2 against 3 forms also a very strong musical opposition (Celibidache 2008).

Finally, for the alignment of chunks in multiple voices, an extra line of dots can visualize the common underlying pulsation, see Table 3.6.

3.3.2 Subdivisions

As proposed in Table 3.5, it is possible to use any form of subdivision in conjunction with the shorthand symbols in Table 3.4. The power of chunking and metric shifts is illustrated by the following examples in Table 3.7. A transcription of Table 3.7 is shown in Fig. 3.3.

Table 3.4 Extended set of shorthand symbols (left) to transcribe chunks of rhythms

ASCII Symbol	Transcription	Binary Form	Category	ASCII code
.		1	unary	46
I		10	binary	73
:		11		58
v		01		118
-		100	ternary	45
<		010		60
w		001		119
X		110		88
>		101		62
+		011		43
i		111		105
H		1000	quarternary	72
II		1010		73 73
::		1111		58 58
.I.		1101		46 73 46
I:		1011		73 58
:I		1110		58 73
-.		1001		45 46
!		1100		33
~	H~I =	100000	tie	126
( )	(-)~I =	00010	silence	40 41
' '' '''	''I '. '.	n/a	appogiatura	39

3.4 Examples

In this section, I present a series of examples to demonstrate the main characteristics of the shorthand notation. They will show that the symbols are easy to memorize and they are quick to write by hand or by using a standard computer keyboard. The notation also helps to reflect the formal structure of the music by helping to visualize musical groupings and formal hierarchies. Bar lines are not necessary because the notation is based upon an underlying small pulsation. It is not dependent on a metrical grid, although bar lines and accentuation marks could be added if needed.

What follows are excerpts from various musical styles including African Percussion, Latin-American Music, Greek Rhythms and Classical Music.

Table 3.5 The main forms of subdivisions. The ratio divides the lengths of the units within the square brackets. If no ratio is given, then the default value of 2 : 1 is applied

Subdivision	Name of the value	Shorthand transcription
1 : 1	Quarter	`I`
2 : 1	Eighth	`[II] or :`
3 : 1	Eighth Triplet	`[3III]`
4 : 1	Sixteenth	`[4IIII] or [::]`
5 : 1	16th Quintuplet	`[5III II]`
6 : 1	16th Sextuplet	`[6III III]`
7 : 1	16th Septuplet	`[7IIII III]`
8 : 1	32nd	`[8IIII IIII] or [[::::]]`
9 : 1	32nd Nonuplet	`[9III III III]`
10 : 1	32nd Decuplet	`[10II III II III]`
11 : 1	32nd Undecuplet	`[11II III III III]`
12 : 1	32nd Duodecuplet	`[12III III III III]`
1 : 4	Whole	`W`
1 : 2	Half	`H`
2 : 6	Dotted Half	`H~I`
2 : 3	Dotted Quarter	`-`
4 : 3	Dotted Eighth	`[4/3I] or [-]`
3 : 4	Half Triplet	`[3/4III]`
3 : 2	Quarter Triplet	`[3/2III]`
5 : 4	Quarter Quintuplet	`[5/4III II]`
5 : 2	Eighth Quintuplet	`[5/2III II]`
7 : 2	Eighth Septuplet	`[7/2IIII III]`

Table 3.6 Alignment of chunks between two voices

Pulse	`.........`
Right hand	`I I I >`
Left hand	`X  X  X`

Table 3.7 The use of different subdivisions together with the pattern I– I: creates a perceptual shift of the pattern in relation to the underlying meter

Subdivision	Shorthand	Time signature
0.5	`[1/2 I- I:]`	9/4
1	`I- I: I- I:`	9/4
2	`[I- I: I- I: I- I:]`	9/4
3	`[3 I- I:]`	3/4
6	`[6 I- I: I- I:]`	3/4
9	`[9 I- I:]`	2/4

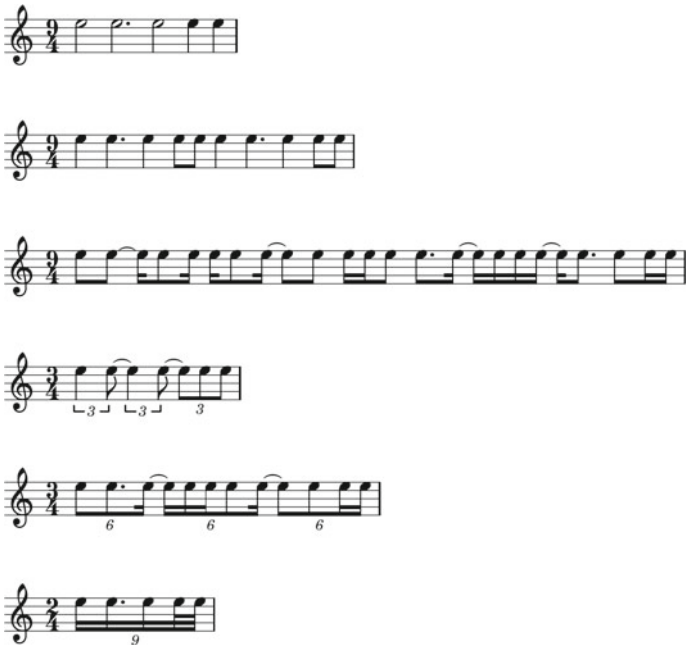



Fig. 3.3 The pattern I– I: subjected to various subdivisions

Table 3.8 Two metric groupings of the same African Ewe pattern: IIXI>

Pulse	. . . . .
Groups	H H H
Pattern	I I .I I >
Pulse	. . . . .
Groups	– – – –
Pattern	> + < >

3.4.1 The Ewe Rhythm

This is the well-known bell rhythm of the Ewe people in shorthand notation: IIXI> Seven beats are distributed over twelve underlying pulses. Musical meter is a concept that does not exist in African drumming (Arom 1991), however, there are at least two ways to layer this rhythm with a regular beat pulsation. Table 3.8 shows the subdivision of the twelve pulses into three groups of four, and into four groups of three pulses. This demonstrates that the shorthand notation can visualize different metric groupings, which has an influence on how one perceives and performs the rhythmic pattern. The latter is very often referred to as *the groove*.

Here is another example of this phenomenon called *metric shifts*. A small variation of the Ewe rhythm: III X> is used in Table 3.9. The number of pulses is twelve.

Table 3.9 Two metric interpretations of a similar pattern: III X>

Pulse	`. . . . . .`
Groups	`H~I  H~I`
Pattern	`I I I X  >`
Pulse	`. . . . . .`
Groups	`I I I I I I`
Pattern	`I I I .I I .`

Table 3.10 Latin-American and African bell patterns and claves

Name	Pulse length	Shorthand
3-2 claves	(334)(24)	`--H IH`
	(343)(33)	`-H- --`
	(3313)(24)	`--! IH`
	(3313)(33)	`--! --`
3-2 Rhumba	(343)(24)	`-H- IH`
Bossa Nova	(334)(33)	`--H --`
Soukous	(334)(15)	`--H :~H`
Gahu	(334)(42)	`--H HI`
Tresillo	332	`--I`
Cinquillo	21212	`>>I`
Baqueto of Danzon	(21212)(2222)	`>>I IIII`
6/8 clave	2212212	`IIX >I`

Beat subdivisions are possible in six, four, three and two. What is shown is the subdivision into two and six. In general terms, any rhythmic pattern over n pulses can be placed into any metric grid of pulsations that is based on the *divisors* of n.

3.4.2 Latin-American Music

Table 3.10 lists a selection of 3-2 claves and bell rhythms used in Latin-American music. The common feature of those 16-pulse rhythms is a characteristic split into $10 + 6$, i.e. three beats are distributed over ten pulses, followed by 2 beats over six pulses. The first half is perceived as carrying impact, or musical tension, whilst the other half brings it to resolution. Samba rhythms behave differently in that they can have a variety of different groupings, see Table 3.11, for example: $4 + 12$ or $7 + 9$. Here the opposition of the ternary beat `> or -` versus the binary beat `I` creates impact, whereas the repetitions of binary beats offer resolution.

Table 3.11 Samba patterns

Pulse length	Shorthand
(22)(322212)	`II -II>I`
(2221)(22212)	`II> II>I`

3.4.3 Greek Verse Rhythms

The analysis of rhythms, especially in in classical music, often refers to the feet of the classical Greek verse in order to explain rhythmic patterns and phrases: Igor Markevitch, Olivier Messiaen and his teacher, Marcel Dupré, all have used the terminology of the Greek verse rhythms in their analytical and pedagogical publications (Dupré 1925; Markevitch 1983; Messiaen 1995). As Georgiades (1982) has pointed out, the classical Greek language and poetry was intrinsically musical because of the measured short and long values of the syllables. When arranged into repetitive patterns, poets were able to create a regular meter, for example the hexameter in Homer's Iliad: `I: I: I: II I: II`,[5] see Silk (2004). Table 3.12 shows the names of the feet, their pulse lengths and their shorthand notation.

3.4.4 Messiaen

Messiaen took inspiration from a great range of different sources: Greek verse rhythms, Indian talas documented in the 13th century,[6] bird songs, and rhythmic palindromes. He published detailed catalogues of his compositional practices that include important accounts on the subject of rhythm (Messiaen 1995). I would like to present an example from his Turangalîla Symphony to illustrate that one can reveal the structure of complex rhythmic passages very easily by using shorthand notation. The final movement uses a counterpoint between the full orchestra, which includes piano and ondes martenot, and one percussionist who uses the woodblocks and a small turkish cymbal. Messiaen often uses bar lines and meters only for the purpose of synchronization. The main theme features frequent metric shifts that break the perception of the ternary 3/16 meter, for example in bars 4-5. Two Trochees in bars 1 and 2 (6 pulses) are followed by an Amphibrach (4 pulses) and a long note (3 pulses), therefore, the 13 pulses of the antecedent phrase exceed the first four bars by one 16th note. The consequent phrase has 14 pulses and splits into eight plus six. The ending of the phrase is a Dactyl with a prolonged final note. This sentence is 27 pulses long. It gets first repeated and then extended at study number 1. The extension features a repetition of the antecedent phrase, so one has two times 13

[5] I130 tòn d'apameibómenos proséphē kreíōn Agamémnōn.

[6] Śārṅgadeva's Saṅgita Ratnākara (Bruhn 2007).

Table 3.12 Greek verse rhythms

Foot	Pulse length	Shorthand
Pyrrhic	11	`..`
Trochee	21	`>`
Iamb	12	`X`
Tribach	111	`i`
Spondee	22	`II`
Dactyl	211	`I:`
Anapest	112	`:I`
Procleusmatic	1111	`::`
Amphibrach	121	`.I.`
Bacchius	122	`XI`
Amphimacer	212	`>I`
Antibacchius	221	`I>`
Peon I	2111	`>:`
Peon II	1211	`X:`
Peon III	1121	`:>`
Peon IV	1112	`:X`
Ionic Major	2211	`II:`
Ionic Minor	1122	`:II`
Molossus	222	`III`
Epitrite I	1222	`XII`
Epitrite II	2122	`>II`
Epitrite III	2212	`I>I`
Epitrite IV	2221	`II>`
Ditrochee	2121	`>>`
Diiamb	1212	`XX`
Choriamb	2112	`I:I`
Antispast	1221	`.II.`
Dochmius	12212	`X>I`
Dispondee	2222	`IIII`
Dactylo-epitrite	2111222	`>: III`

pulses. Again, it is the binary Amphibrach `.I.` that disturbs the regular ternary meter. Adding to this metric ambiguity between binary and ternary chunks are two accents on the first *and* on the last 16th note of the Amphibrach. A rhythmic counterpoint to these sentences is played by woodblock and small turkish cymbal in unison: `H H H I - I I I I - - - X - H HH`. Because its length of 52 pulses is not an integer multiple of 27, there are ongoing metric shifts between the orchestra's sentences against this repeating counterpoint. With its many chunks of 4 pulses and 2 pulses, the counterpoint contrasts the predominantly ternary rhythm of the full

orchestra. With the orchestra in the forefront, the percussive counterpoint creates a polymetric situation and it also works very similar to a hocket, because it frequently falls in-between the notes of the orchestra.

```
$A bars 1 - 9
antecedent consequent
>>.I.-     :ii:I~I
$A bars 10 - 18
>>.I.- :ii:I~I
$B num 1;
>>.I.>>>.I.-:ii:i.:i.:I~I
$A
>>.I.- :ii:I~I
$C
>>.I.- :i: iiii iii:I~I
```

3.4.5 Beethoven

In Beethoven's music one admires his capacity to use rhythm as a means for building thematic contrasts and energetic movement. In his analysis of Beethoven's Symphonies, Markevitch (1983) uses Greek verse meters as a system for analysis and he refers also to the articulation and placement of accents in German language. In order to demonstrate the flexibility of shorthand notation, here is a transcription of the theme of the 32 Variations in c-minor, WoO 80. This is not just a display of the rhythm of a Sarabande `I-.`, but through the use of various forms of articulations, silences and anacrusis it is clearly driving forwards until the expansion of the theme reaches its maximum in bar six, and, after a rhetoric silence, the tension is rapidly released in a simple cadential downward movement to the tonic.

```
$bar 1
I~I I~I~-.
$bar 2
I(I) I~I~I [5IIIII]
$bar 3
I(I) I~I~-.
$bar 4
I(I) I~I~I [::]
$bar 5
I(I) I~I~I [::]
$bar 6
I(I) I~I~I~I
$bar 7
(-). I(I) I(I)
$bar 8
I(I) I(I) I(I)
```

3.4.6 Mussorgsky

Mussorgsky's opening *Promenade* from his *Pictures of an Exhibition* is a masterpiece of economy - just two different durations, quarters and eighth notes, are used to illustrate a walk towards and between the pieces of the exhibition. Interestingly, one finds mirror-like melodic movements, opening and closing phrases, in the first two bars, a change of meter between 5/4 and 6/4, and over the course of the piece there is a wonderful variation of the dactyl motive, which is presented twice at the beginning, `:I:I`. Afterwards, it is developed and varied to become `::I` until the dactyl figure spans an entire bar of six quarters in bars 9 and 13, `:::::I`. After this climax, the expansion of the dactyl is reversed again to become `::I` and `:I`. The movement of expansion and contraction, of breathing in and out, of accelerating and slowing down, of experiencing impact and resolution, these phenomena are not only characteristic for music of the Romantic period, but they are much more general and frequently used across different styles and diverse musical cultures (Giger 1993). They are clearly manifest on the level of rhythm and through the use of rhythmic patterns, phrases and sentences. The shorthand notation provides a good framework for visualizing the musically very important forces of tension and release.

```
$Promenade, bars 1 and 2
III:I :IIIII
$ bars 3 and 4
III:I :IIIII
$ bars 5 and 6
III:I :III:I
$ bars 7 and 8
III:I :III:I
$ bars 9 and 10
III::I :::::I
$ bars 11 and 12
III:: III::I
$ bars 13 and 14
:::::I :IIIII
$ bars 15 and 16
:III:I :II:I:
$ bars 17 and 18
IIII:: III:II
$ bars 19 and 20
I::III :IIIII
$ bars 21 and 22
:IIIII I:I:II
$ bars 23 and 24
IIIIII :IIIII
```

Table 3.13 Compound polyrhythm of 2:4:9:3 in shorthand notation *s*

2	↓						↓					
4	↓			↓			↓			↓		
s	H	H	.	–	H	I	I	H	–	.	H	H
9	↑	↑	↑		↑	↑		↑	↑		↑	↑
3	↑				↑				↑			

3.4.7 Debussy

Debussy's beginning of Syrinx is another example of a climax technique that involves great rhythmic contrasts. This time it is achieved through the use of dotted rhythms versus regular sixteenth notes and versus sixteenth note triplets, which creates a special contrast of 2:3. A dramatic silence is used to emphasize the climax, which presents the longest note duration of this section.

```
$Syrinx beginning of
-[:]-[:]:: -[:] I~I~I~I
$ bar 3
-[:]-[:]:: -[:] I~[3III] [3III][3I'II]
$ bar 5
-[:]I~I [3(I)II][3III] (-). I~I~I~I~I~I~I~I I~I
$ bar 8
I~I~I~I I~I
```

3.4.8 Polyrhythm

This section will demonstrate that it is easy to write polyrhythms with the proposed shorthand notation and that it is a good tool in order to learn how to execute polyrhythms as well. Consider the polyrhythmic pattern of 2:4:3:9. over thirty-six pulses. The sequence of pulses per beat is: 4, 4, 1, 3, 4, 2, 2, 4, 3, 1, 4, 4. Table 3.13 shows how one can write it as shorthand. Small arrows indicate the layers of 2, 4, 9 and 3 beats, which together form a compound pattern *s*.

3.4.9 Conclusion of Examples

The examples have shown that the proposed system of shorthand notation is able to capture music of varying rhythmic complexity and from different periods and cultures. It can do this because of the basic psychological principle of chunking, and because it is not dependant on the concept of meter and and its associated

patterns of accentuation. Therefore, and this has important musical consequences, the visualization of the difference between binary and ternary chunks of rhythm is very easy to realize. It is then easier to analyze and to musically execute the alternation of the impact of a certain rhythmic pattern and its subsequent resolution, as shown in the examples of African and Latin-American bell patterns and claves. We have also seen that the rhythm of the Greek Verse remains a very helpful system to explain prominent rhythmic structures in music of different styles, for example in Beethoven, Mussorgsky and Messiaen. The shorthand notation makes it easy to quickly grasp an overview of the rhythm and form of a piece of music, and it is helpful in discovering the build-up of a climax when it is played out by contrasting patterns and through the use of different note lengths, see Debussy's Syrinx and Beethoven's c-minor theme. Another interesting feature provided by the system is the aid to capture and to execute polyrhythms and metric shifts. The examples of using subdivisions in combination with basic chunks, as well as the demonstration of different metric interpretations of the same rhythm pattern, are able to support this point.

3.5 Conclusion

The proposed shorthand notation has a number of advantages: It is fast to write and to read, and easy to recall due to the cognitive principle of chunking. It is intuitive, for example the symbol 'X', which represents two beats and a silence, is made of two pen strokes. The shorthand notation can be extended to cover all cases of Western music notation, for example subdivisions (tuplets), polyrhythms and metric shifts. The notation is not making assumptions about any fixed metric structure, the only condition is an underlying common pulsation. The electronic storage of the notation can be accomplished through ASCII characters in text files or databases It is easy to parse the format and to convert it into a different output format, for example *lilypond*, MIDI, or MusicXML. Database entry and searches are easy to manage, and the proposed notations support these powerful tools for music analysis. The many music examples from different eras and different styles underpin the general purpose and usage of this particular method of notation. The shorthand notation also lends itself to the area of Ethnomusicology where the standard notation with its implied metric structure is seen as problematic (Arom 1991).

The dualism between meter and rhythm enables a technique, in which rhythmic phrases are shifted away or towards metric beats. If the length of the pulse sequence that forms the rhythmic phrase is a highly divisible number, then the same phrase may be re-interpreted in alignment with different forms of meters where the length of a beat is a divisor of the number of pulses in that phrase. A change of the length of the beat leads to a shift of metric accents and to a change in the perception of tempo, which is a very interesting compositional device.

Equipped with this new way of notation, we can proceed to make use of it in the following chapters, which deal with rhythm analysis and composition.

References

Arom, S. (1991). *African polyphony and polyrhythm*. Cambridge: Cambridge University Press. (ISBN 052124160X).

Barlow, C. (2008). *On musiquantics*. Musicological Institute, Musikinformatik und Medientechnik of the University of Mainz. (ISSN 0941-0309).

Bruhn, S. (2007). *Messiaen's contemplations of covenant and incarnation: Musical symbols of faith in the two great piano cycles of the 1940s*. Hillsdale, NY: Pendragon Press.

Celibidache, S. (2008). *Über Musikalische Phänomenologie*. Augsburg: Wissner-Verlag. (ISBN 978-3-89639-641-9).

Chew, G., & Rastall, R. (2001). *The new Grove dictionary of music and musicians* (2nd ed., Vol. 18, pp. 140–189). New York: Macmillan Publishers Ltd. (S. Sadie (Ed.). ISBN 1-56159-239-0).

Demaine, E. D., GomezMartin, F., Meijer, H., Rappaport, D., Taslakian, P., Toussaint, G. T., et al. (2009). The distance geometry of music. *Computational Geometry*, *42*(5), 429–454. https://doi.org/10.1016/j.comgeo.2008.04.005.

Dupré, M. (1925). *Traité d'Improvisation à l'Orgue* (Vol. 2). Paris: Alphonse Leduc. http://quod.lib.umich.edu/. Retrieved December 2010.

Georgiades, T. G. (1982). *Music and language: The rise of Western music as exemplied in settings of the mass*. New York, Cambridge: Cambridge University Press.

Giger, P. (1993). *Die Kunst des Rhythmus: Professionelles Know How in Theorie und Praxis*. Mainz, London, New York: Schott. (ISBN 978-3-7957-1862-6).

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*(1), 1–40. https://doi.org/10.1006/cogp.1996.0011.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236–243. https://doi.org/10.1016/S1364-6613(00)01662-4.

London, J. (2012). *Hearing in time. Psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Markevitch, I. (1983). *Die Sinfonien von Ludwig van Beethoven. Historische, analytische und praktische Studien*. Leipzig, GDR: Edition Peters.

Messiaen, O. (1995). *Traité de Rhythme, de Couleur et d'Ornithologie* (Vol. 1). Paris: Alphonse Leduc. (ISBN 978-004628892).

Miller, G. H. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Sethares, W. A. (2007). *Rhythm and transforms*. Berlin: Springer. (ISBN 9781846286391).

Silk, M. S. (2004). *Homer, the Iliad*. Cambridge, New York: Cambridge University Press.

Toussaint, G. (2004). A comparison of rhythmic similarity measures. Technical report. School of Computer Science, McGill University. http://cgm.cs.mcgill.ca/~godfried/publications/similarity.pdf. SOCS-TR-2004.6. Retrieved March 2018.

Chapter 4
Partitions and Musical Sentences

4.1 Introduction

This chapter presents a model for asymmetric rhythmic sentences on the basis of integer partitions and combinatorics. It has been implemented into *chunking*'s *sentence* method.[1] The model uses a technique, which subdivides a series of pulses into smaller and smaller sections until what is left are so-called rhythmic chunks. These are the binary or ternary cells that are found in Greek verse rhythms, for example, and which are represented by symbols in SNMR (see Chap. 3, Sect. 3.4.3). Through certain combinations of binary and ternary chunks one arrives very quickly at an asymmetric order. This is a key concept that can also be found in non-isochronous meters. For example, one of the Turkish Aksaks `I:I >`, has the beat lengths $\{2, 2, 2, 3\}_9$. Aksak[2] rhythms are a family of rhythms, which have been studied in detail by Arom (1991) and Brăiloiu (1951). Combinatorial concepts are also integral to the rich tradition of North- and South-Indian classical music (Junius 1983; Sankaran 2010), where a long sequence can be partitioned asymmetrically with respect to the underlying beat, in order to create polyrhythmic climaxes that are released by coinciding again with the first beat of the tala. The joy of this return to the first beat is often visible in the gestures and body-languages of performers and audiences.[3] The reader might also be interested in the research by Rafael Reina on Karnatic music (Reina 2014), which is supported by many sound examples.

Asymmetry plays an important role also in Western music of the late 19th and 20th century, notably in the music by Igor Stravinsky and in the works by Olivier Messiaen. This chapter will describe the asymmetry that results from a hierarchy within the creation of musical form, which develops upwards from chunks into patterns, into phrases, and finally into complete sentences and sections of a musical piece. The metric grouping of phrases into two or three pulses is also investigated,

[1] https://github.com/gboenn/chunking.

[2] The term Aksak is Turkish and means 'limping'.

[3] See, for example, online available video recordings by master performers such as Trichy Sankaran.

and a computational measure for metric complexity is introduced. The next section begins with a brief description of the mathematics of integer partitions, then I will discuss the essential structure of a musical sentence, followed by an analysis of typical examples found in the music by Stravinsky and Messiaen. Finally, the argument made for the application of integer partitions will be supported by insights from the field of cognitive psychology.

4.2 Integer Partitions

Every natural integer n has a finite number of partitions $p(n)$. Every n can be written as a unique sum of smaller integers. Partitions are multisets of integers whose sums result in n. We notate partitions as (multi-)sets of integers with subscript n, for example, $\{7, 5, 4\}_{16}$ is equivalent to writing $16 = 7 + 5 + 4$. The order within the partition is not relevant, which means that the sequence of terms can be changed, but a permutation would always be counted as the same partition of n. For example, the partitions of 5 are shown in Table 4.1. When used as a rhythm, then the parts define the time elapsing between beats (or note onsets) measured in number of pulses. They can be easily written in shorthand notation. The terms in the partitions b to f can be re-arranged, hence they can be used to generate different rhythmic structures.

In addition, there are two possible interpretations of partition f. It becomes sometimes necessary to decide whether to choose a binary or a ternary chunk. Of course, this would decide over the metric grouping and, therefore, the musical context has always to be taken into account.

The number of permutations $s(p)$ that can be generated for each partition is dependant on the number of parts k *and* on the number of distinct parts. Counting the repetitions of distinct parts as m_1, m_2 and so forth, with $m_1 + m_2 + m_3 + \cdots + m_r = k$, the number of permutations is given by

Table 4.1 All partitions of 5 and their equivalent rhythm in shorthand notation

Partitions of 5						
	Terms					Transcription
a	5					`-~I`
b	4	1				`I~I.`
c	3	2				`-I`
d	3	1	1			`-:`
e	2	2	1			`I>`
f	2	1	1	1		`>:` or `Ii`
g	1	1	1	1	1	`i:` or `:i`

$$s(p) = \frac{k!}{m_1!m_2!m_3!...m_r!} \tag{4.1}$$

For example, partition f in Table 4.1 has $\frac{4!}{1!3!} = \frac{24}{6} = 4$ permutations: >: X: :> iI .[4] A special case are partitions into k distinct parts without any repetition. Their number of permutations is simply $k!$.

A few statistical descriptors are useful for the analysis of integer partitions: The mean is n/k. The standard deviation σ gauges the spread of values within a certain partition, whereas the range denotes the difference between the maximum and minimum integer used as a part. We define the beat density of a partition as $D = \frac{k}{n}$. The parts themselves define the time elapsing between note onsets, which is measured in number of pulses.

4.2.1 Partitions into k Distinct Parts

Partitions of n in k distinct parts are important cases, for example $\{4, 3, 2, 1\}_{10}$, which is the famous *tetraktys* of the Pythagoreans. In this chapter, we are putting emphasis on this class of partition, because it leads to a great variety of musical rhythms. The number of partitions of n into $k > 1$ distinct parts equals the number of partitions of n into odd parts. In Table 4.1, the partitions a, b and c have distinct parts, the partitions a, d, and g have odd parts only.

The partitions of n in k distinct parts are interesting because as a structure of a rhythmic phrase they easily create tension in relation to an underlying regular beat or pulse. For example, instead of playing along with the beat, such as $\{2, 2, 2, 2\}_8$, we could play against it in an asymmetric way, $\{3, 5\}_8$, or, after partitioning the 5: $\{3, 3, 2\}_8$. The discrepancy between a rhythmic pattern and an underlying even beat structure creates syncopation and thereby musical tension.

There are certain lengths of phrases n, which will allow for an easy change of the underlying metric structure. The perception of the rhythm will change too, because of the different metric feel. For example, let us consider a rhythmic phrase of length $n = 30$. The *divisors* of 30 are: $\{15, 10, 6, 5, 3, 2\}$. Therefore, for any the rhythmic phrase of 30 pulses, it is easy to arrive at six different metric variations. See, for example, Table 3.7 and Fig. 3.3 in Chap. 3. Here, meter is understood as a cyclic structure with equidistant beats. The beats form an integer ratio with the underlying pulsation. Rhythmic patterns are always aligned to a metric structure. A rhythmic phrase structure that spans over a highly divisible number of pulses (such as 30, or 60 for example) can be realigned to metric structures that use one of the divisors of that number as a constant distance between the beats. The number of divisors equals the number of metric variations. Some numbers are highly divisible, others have relatively few divisors, and the prime numbers have none.

[4] This is also an important aspect of the Burrows-Wheeler Transform, which will be discussed in Chap. 5.

4.2.2 *Partitions into Parts with an Arithmetic Progression*

The partition of n into k distinct parts, where the parts form an arithmetic progression with $\Delta = 1$ is the unique class of partition where all parts are very close to the mean, i.e. the standard deviation of this form of partition is closest to zero when compared with all other partitions into distinct parts. For example, given: $\{7, 6, 5\}_{18}$, then $\sigma \approx 0.816497$ reaches its smallest value amongst all partitions of 18 with 3 distinct parts. Of course, for other integers, Δ and σ can be different. The generating algorithm for such partitions has been studied by Bouroubi and Tani (2009).

I propose that the following scenario of partitions of a number of pulses into distinct parts generate contrast or impulse in musical terms if the following condition is met: *Subsequent chunks of rhythm patterns should have lengths that are coprime.* This condition is met if the parts have arithmetic progressions with $\Delta = 1$, because consecutive integers are always coprime. The two most important elementary chunks are the first prime numbers 2 and 3. Table 4.2 shows a few examples of interesting grooves that are based on this principle. An exception to the rule is that chunks of the same size may be repeated.

Uneven beat structures are not uncommon in music. The traditional theory of meter is based on an evenly spaced grid of pulsations with either binary or ternary subdivisions of the beat (London 2012). However, Polak (2010) reported the practice of non-isochronous beat subdivisions in the jembe music of Mali. Rainer Polak found a pulsation grid that is not evenly spaced, where the pulses themselves are irregular and not equidistant. He questioned the view that an isochronous underlying pulsation is a 'universal' principle (Polak 2015). Therefore, in some musical cultures, the micro-timing structures of performances cannot be mapped to a common equidistant pulsation.

Number theory is often concerned with the problem of *counting* the number of partitions $p(n)$ of a certain kind. This is accomplished by so-called generating functions, power series in particular, see Hardy and Wright (2008). Of further interest are also relations between different kinds of partitions. For example, the number of partitions of n whose largest part is k is equal to the number of partitions of n with k parts. Or, the number of partitions of n with no parts equal to 1 is $p(n) - p(n-1)$ (Wilf 2000).

Knuth (2013, p. 408) published an interesting algorithm for generating all partitions of n by using a binary tree.

4.3 Musical Sentences

In Chap. 2, we stated that there is a causal continuity in music. Let us investigate in more detail the cause and effects, which lead to this continuity. The main cause for continuation in music is the insertion of musical impulse, of musical energy. It can be created in a number of ways: If we leave out changes in pitch and concentrate just on

Table 4.2 Examples of grooves where chunk sizes are coprime

	Partition of n	Example
a	$\{2, 3, 2\}_7$	`I-I`
b	$\{2, 3, 2, 2\}_9$	`I-:I`
c	$\{2, 3, 2, 2\}_9$	`I-I:`
d	$\{2, 2, 3, 2\}_9$	`II-:`
e	$\{2, 3, 3, 2\}_{10}$	`I--I`
f	$\{3, 2, 2, 3\}_{10}$	`XII-`
g	$\{3, 2, 3, 3\}_{11}$	`-I--`
h	$\{2, 3, 2, 2, 3\}_{12}$	`I-II-`
i	$\{2, 3, 3, 2, 2\}_{12}$	`I--II`
k	$\{2, 3, 3, 2, 2\}_{12}$	`I->II`
l	$\{3, 2, 3, 2, 2\}_{12}$	`XIXII`
m	$\{2, 3, 2, 3, 3\}_{13}$	`I-I--`
n	$\{3, 2, 2, 3, 3\}_{13}$	`XII->`
o	$\{2, 3, 2, 3, 3\}_{13}$	`I-I+>`
p	$\{2, 3, 3, 3, 3\}_{14}$	`I--X>`
q	$\{2, 3, 2, 2, 3, 2\}_{14}$	`I-I:-I`
r	$\{3, 2, 3, 3, 2, 2\}_{15}$	`XI--II`
s	$\{3, 2, 2, 3, 2, 3\}_{15}$	`XIIXI-`
t	$\{2, 3, 2, 3, 3, 3\}_{16}$	`I-I-->`
u	$\{3, 2, 3, 2, 3, 2, 2\}_{17}$	`XIXIXII`
v	$\{2, 3, 2, 3, 3, 2, 2\}_{17}$	`I-I--II`
w	$\{2, 3, 2, 3, 3, 2, 3\}_{18}$	`I-I--I-`
x	$\{2, 3, 2, 2, 3, 2, 3, 2\}_{19}$	`I-IIXIXI`

one and the same tone, then impulse can be generated by repetition, by speeding up repetition, by changing the dynamics of the tone from soft to loud (crescendo), and by changing the articulation of the tone from long to short. In this account there are two parameters that affect the timing of the note: First, the time elapsing between two notes, i.e. the occurrence of a sound in time, and secondly, the absolute duration of the note, which is the length of the sound. Both factors combined constitute musical rhythm. In general, a rapid succession of short notes generates impulse, whereas less note events and longer durations are associated with calming down, and with a dissipation of musical energy (Thakar 2011).

This leads us to two more questions: What is a short and what is a long duration? And, is there a way to define the density of musical events? Let us bear in mind that these two parameters, duration and density, together generate musical form by establishing rhythmic movement. And, although one can find absolute timing values that form psychological limits within which musical rhythm and meter exists, one must admit that there is no short or long duration in absolute terms outside of

the musical context, because there is no absolute value of a musical tempo, which exists outside of the reality of the performance. Tempo is a very important condition under which listeners can have a unified musical experience (Fischer et al. 1986a, b; Celibidache 2008). Therefore, it is much better to talk about the relations between musical events. It is more relevant to discuss the proportions and relative lengths of musical rhythm. There are certain proportions that generate musical impulse whereas others are more conducive to bringing the energy level down, to lead to an ending or a conclusion. The exploration of various kinds of rhythmic movement is the task of the composer, or in case of improvisation, it is an integral part of the performance.

When we discussed organic form we saw that musical form is based on growth, balance and decline. The reason that there is no indefinite, unlimited growth is because decline and death follows in all natural processes. We also introduced the image schema, which reflects this process, and that is the *musical triangle*. A piece of music is an articulated dynamic process. It means that the stages of growth, balance and decline define the whole entity of a composition, and they permeate into the lower stages and sections of the piece. The articulation of the triangle can cause even the smallest parts of a piece to follow the same dynamic process.

In the following sections, after going through the analysis of some musical examples, I will develop an algorithm for the composition of a rhythmic sentence, which will be defined as an entity in which impulses and resolutions are played out, and which consists of two or more phrases, each one composed out of one or more patterns, which break down further into chunks.

There is one aspect of musical meter that is also linked to this algorithm. Non-isochronous meters (NI meters) are meters that have different beat lengths, which alternate in very specific patterns. They have been analyzed by London (London 2012) who defines them with a series of well-formedness rules. With *chunking*'s *sentence* algorithm it is possible to generate NI meters. It is even possible to construct sequences of changing NI meters, which are held together by an over-arching musical concept, namely tension and release. Further inspiration was drawn from the analysis of Eastern-European rhythmics, notably the Aksak rhythms (Arom 1991; Baraldi et al. 2015; Brăiloiu 1951). The theory of Aksak offers prime examples of NI meters. An Aksak meter can be described as a multi-set of beats that are either 2 or 3 pulses long, for example $\{2, 2, 2, 3\}_9$.

Furthermore, the theories on NI meter and Aksak rhythms led me to develop a measurement of the complexity of such metres and rhythmic phrases in general. The main idea for this measurement is that a change from a group of 2 to a group of 3 pulses generates some form of musical energy or interest that would not been there if the metre or phrase would just consist of either binary or ternary groups of pulses. Furthermore, it is conceivable to have much longer groups of 4, 5, 6, and more pulses. The crucial aspect of the notion of complexity is the coprimality of the groups that are immediately following each other, for example the phrase $\{2, 3, 2\}_7$ has two transitions: One going from binary to ternary and the other going back from ternary to binary. It is easy to understand why the transition between two coprime numbers, 2 and 3, generates musical energy. It is simply because 2 does not subdivide 3 without remainder. There is a difference, however, when the integers involved are

not coprime, for example $\{2, 4, 2, 6\}_{14}$. Here, the larger groups are integer multiples of 2, and, as a result, there is much less metric conflict, if any, compared to the previous example. Therefore, the understanding of the prime-number composition of metric groups and the comparison of neighbouring groups can lead to a measure of metric complexity. The details of this algorithm will be given in Sect. 4.5.

Miller (1956) made an important discovery in the field of cognitive psychology by finding a 'magic number' that reappeared again and again in various experiments to test the capacity of human cognition and memory. The number 7 ± 2 seems to accurately describe how many items can be held in short-term memory. This number can be increased by a technique called 'chunking'. If a long list of items is cognitively separated into groups (chunks) it will be easier for the mind to grasp, memorize and to reproduce the entire list correctly. Gobet (1996) and others have expanded Miller's theory. Their model includes a template mechanism, which allows for slots with variable content in addition to a core of chunks. A chess position, for example, is memorized by building chunks of pieces in the mind, which group together via proximity. There is a certain number of fixed chunks (groups of chess pieces), but also slots which may contain a different variation of a chunk. In this case only a pointer is needed that will lead to the right chunk in long term memory. Chunks and templates facilitate information processing in our minds. The *sentence* algorithm that we will describe in this chapter, Sect. 4.7, makes use of both concepts for the creative process of composing rhythmic phrases.

4.4 Asymmetric Sentences

In this section, I would like to show how asymmetric musical sentences are based on the principle of coprimality of metric groups. We will see changes in metric groups, for example, the partition $\{2, 3, 5, 4\}_{14}$, which contains a sequence of coprime integers. We will also look at changes of rhythmic density as expressed by note onsets per bar.

4.4.1 Stravinsky's Game with Metric Asymmetry

In Igor Stravinsky's *Histoire du Soldat*, the 1st movement is the soldier's march. It is a parody of a military march. There are two strategies that create asymmetry (the following transcriptions use the mapping I = ♪). First, phrases are started and broken off: `-. -. IIII`. The second strategy consists of the insertion of a ternary group `III` to throw the listener out of step with the marching ostinato played by the double-bass: `IIII IHI III HH`.

In his work *Symphonies of Wind Instruments*,[5] Stravinsky uses two more strategies for asymmetry: Coprimality of metric groups, and repetition. Coprimality is also found sometimes in note density per bar. The bassoon theme at rehearsal number 8 has the following rhythm and metric grouping: `:I i II[:] X >.[:] X >I`, in pulses: {4, 3, 5, 3, 5, 3, 5}. The following grouping of the theme reveals the grouping on the lowest level (chunks of 2 or 3): `: I : X I[:] X > .[:] X > I`, and in pulses: {2, 2, 2, 3, 3, 3, 3, 2, 3, 3, 2}. This theme is placed against the backdrop of a 2/4-meter, and the three flutes emphasize the polyrhythmic structure of the bassoon theme by using a 2:3 proportion, i.e. `I against -`. The sequence of note onsets per bar is for the bassoon theme: {3, 4, 3, 3, 4, 3, 2}. The principle of coprimality governs the metric grouping of the theme on two levels: chunks and patterns. It also seems to control the note density per bar for the instrument that presents the theme. If the onsets of all four instruments are taken into account,[6] then the following onset density is revealed: {3, 4, 5, 4, 5, 4, 3}. Again, coprimality. Could this be an accident? It gets even better: If we count each individual instrument's onset, summing them together when the onsets fall onto the same time, the following picture emerges: {4, 1, 4|4, 1, 4, 1|1, 3, 1, 1, 1|4, 1, 1, 3|1, 1, 4, 1, 1|4, 1, 1, 3|1, 1, 4}. The onset density per bar is {9, 10, 7, 9, 8, 9, 6}, where, except for the last bar, we find that consecutive values are coprime.

The coprimality principle is also demonstrated at the beginning of the piece. The trumpets accentuate the metric grouping: `I-I-I-~I`. But, together with the trombones, the following compound patterns arises: `:-IXII-`. The entire first sentence is 17 pulses long, and can be divided into two phrases of 7 and 10, or written as partition $\{7, 10\}_{17}$. The woodwinds present the melody with the following rhythm: `IIX I::H`. Note also the overlapping succession of the 7-pulse pattern `IXI` between the woodwinds and the compound brass rhythm.

It is characteristic for Stravinsky's style that there can be an ambivalence between the metric grouping that is suggested by the notation, and the layer of rhythmic patterns, so that the actual sounding metric grouping dominates over the notated grouping. Specifically, the bars 3 - 5, have the notated grouping `-I-~I`, but the syncopation, embellishments, and the melodic profile persuade us to hear a syncopated 3/4-meter: `I:: H`, or with embellishments: `I:''.''. H`.

The following chordal sentence between rehearsal number 1 and 2, has two phrases: $\{19, 14\}_{33}$, which break down into the following patterns: A contraction of patterns for the first phrase: `(I)H~: (I)-~: (.)-`, or $\{8, 7, 4\}_{19}$. Then, a new idea `IIII`, and a reduced chord pattern `(.)H(.)`, or $\{8, 5, 1\}_{14}$ together build the second phrase. We find again coprimality of the phrase lengths, and of the pattern lengths, combined with the organic principles of repetition and contraction.

Finally, Stravinsky's ballet *The Rite of Spring* is a celebration of metric asymmetry, of coprime metric groups clashing violently, of the dancing ecstasies of 2s and 3s. And there is at times mechanic, brutalist repetition. Messiaen found in this work his "personnages", which are rhythmic themes that act like characters on a stage.

---

[5] 1947, revised version.

[6] Onsets that fall on the same time within the bar count as one onset.

All of these works provide many evidences for the conscious use of coprimality as a tool for creating musical impulse.

4.4.2 Messiaen: The Birds as Teachers of Composition

Olivier Messiaen's music would be unthinkable without his ornithological studies and transcriptions of bird song. Nearly every work quotes singing birds in one way or the other. From the 1950s onwards, until the end of his life, he transcribed bird songs systematically from all around the world. In 1960, he finished a major orchestral work called *Chronochromie*, which was first performed at the Donaueschingen festival (Messiaen 1963). There are seven movements: *Introduction, Strophe I, Antistrophe I, Strophe II, Antistrophe II, Épôde* and *Coda*. The movement *Épôde* is an 18-voice polyphony of solo strings that is intended to be the artistic rendering of a morning choir of birds at early sunrise. The orchestration of the entire work is large and contains difficult solo parts for the mallet instruments in the percussion section, but it is without solo piano. In the movements *Antistrophe I* and *Antistrophe II*, two bird songs are alternating: The song thrush (grive musicienne) and the Eurasian skylark (alouette des champs). The tempo indication is always *un peu vif*, with ♪= 160 for the song thrush sections and ♪= 144 for the skylark. The instrumentation uses a large woodwind section for the song thrush, and it uses mallet instruments (Glockenspiel, Xylophone, Marimba and Tubular Bells) for the skylark, which are placed before a background of sound dust generated by the higher strings. If we call the song thrush music A, and the skylark music B, then the sections have the following form. *Antistrophe I*: A, B, A', B', A". *Antistrophe II*: A"', B", A"", B"', A""'. Over the course of both movements, both bird songs become longer and increasingly rich in rhythmic variation. We would like to demonstrate this in the following analysis, where the transcriptions into SNMR use the mapping .

This analysis will demonstrate the strategy of growth. Although being very different, the song thrush and the skylark songs have some simple rhythmic patterns in common. There are 22 occurrences of the Anapest pattern `:I`, which start a new musical phrase in both bird songs. It happens nine times in *Antistrophe I* and thirteen times in *Antistrophe II*. In both movements, the Anapest pattern is a seed pattern for growing longer phrases over the course of each section of bird song. The song thrush in *Antistrophe I* uses the Anapest for the first time in the following phrase: `:I(I)::I(I)(HH)`. Later in the movement, the Anapest leads to `:IH`, and `:IHH`. The skylark song also features growing phrases of increasing complexity, all of which starting in the same way with an Anapest: `:I`, `:IiI`, `:I:IiI`, `:I:Ii:IiiI`. Here one can nicely observe how the ternary chunk `i` introduces musical impulse via the metric opposition of groups of 2 and 3 pulses. Quite often during *Antistrophe II*, the skylark uses chains of two or three Anapests: `:I:I`, `:I:I:I`.

In *Antistrophe II*, one finds a similar growth of the pattern `iI`, which is the Greek verse meter Peon IV. Among the phrases of the song thrush we find, in the order of appearance:

`iIiI`,

`iI(I)iI(I)`,

`iI(I)iI(I)iI(I)iI(I)iI(I)(II)`,

`iI(I)iI(I)iI(I)(II)`.

The skylark uses the Peon IV to build quite different phrases: `iI::I`, `iI`, `iIIiI`, `iIHH`, `iI'I'I'IHH`, and finally `iIiI:I`.

So far, we have analyzed phrases, which start with a common pattern and evolve subsequently. Of course, we can also observe variable growth on the left side of the pattern. Based again on `:I`, we find expansions like `::I::I::I`, or the even more complex `I:i:I::HH`. Often at the end of a phrase, the bird songs double `I` into `H` like a punctuation mark. During both movements, the skylark uses rhythmic variations, which become more and more intense as it ascends higher and higher into the sky. But, also the song thrush is a master of variation. Interestingly, variation is paired with repetition. As we observed before, repetition in its essence is the simplest form of variation. Because of the causal continuity in music, what has sounded before alters the perception of the present sounds. If repetition is then also paired with increase in loudness, for example by using crescendi and accents, then we have found one of the basic models of building climaxes, for example in the last occurrence of the song thrush in *Antistrophe II*, see Table 4.3.

The principle of intensification by repetition can be found in both bird songs. However, the Greek Trochee, `>`, is characteristic only for the song thrush music. Already in *Antistrophe I*, we found the phrases: `(II)>II`, `(I)>II`, `>I(I)`, and `>H>H`. `>II` is the Greek Epitrite II, `>I` is the Amphimacer. These phrases have a distinct signalling character.

The Diiamb `XX` is prominent in the skylark phrases. In *Antistrophe I* we find `XX`, and `XX::IiiiI`. In *Antistrophe II* it is used three times in this form `XXXI`, which could be viewed as a combination of Diiamb with a Bacchius `XI`. The Diiamb is varied into the first Epitrite, `X:I`, and expanded into `XXX::I`.

The Tribach `i` is used in abundance in order to begin phrases in both bird songs. Table 4.4 show that a great variety of rhythms, especially in *Antistrophe II*, is produced from this seed pattern. In his *treatise*, Tome 5, Vol. 1, Messiaen himself analyses pages 111-113 of the score, to talk about how he used the song thrush - one of his favourite birds - in *Chronochromie*. Although discussing mainly his treatment

Table 4.3 Song thrush (Grive Musicienne), 3rd song in Antistrophe II, p.125 of the study score

Phrases	Length	Onsets
::iI(I)::iI(I)	22	16
>H>H	14	6
iI(I)iI(I)iI(I)iI(I)iI(I)(II)	39	20
(I)>II	9	4
(I)>II(I)	11	4
iI(I)iI(I)iI(I)(II)	25	12
:I:I:I:I	16	12
vi'H	9	5
vi'H(I)	11	5
I:I(I)I:I(I)I:I(I)	24	12
i:I(II)i:I(II)	22	12
>I(I)>I(I)	14	6
II(II)II(II)II(I)	22	6
i:II(I)i:II(I)	22	14
::iI(I)::iI(I)	22	16
::I(I)	8	5

Table 4.4 Variations of the Peon IV pattern iI found in the phrases by the song thrush. The second column shows the lengths in pulses. The third column is the number of onsets

Phrases	Length	Onsets
iiHH	14	8
iiH	10	7
iHiHiH	21	12
i:I	7	6
iIiI	10	8
iiI(I)	10	7
iI(I)iI(I)	14	8
iI(I)iI(I)iI(I)iI(I)iI(I)(II)	39	20
iI(I)iI(I)iI(I)(II)	25	12
i:I(II)i:I(II)	22	12
i:II(I)i:II(I)	22	14

of harmony, color and instrumentation, he singles out the Peon IV iI at rehearsal number 67.

In Table 4.5, I have listed the skylark variations of Peon IV and all phrases starting with the Tribach. These are found in the skylark sections of *Antistrophe II*.

Table 4.5 Variations of the Peon IV pattern iI found in the phrases by the skylark. The second column shows the lengths in pulses. The third column is the number of onsets

Phrases	Length	Onsets
`iHiHiHiH`	28	16
`iI::I`	11	9
`i:iI`	10	9
`iI`	5	4
`iIIiI`	12	9
`i>I`	8	6
`iiiH`	13	10
`iHHH`	15	6
`iIHH`	13	6
`iI'I'I'IHH`	19	9 (12)
`iiiiI`	14	13
`i:I`	7	6
`iiI`	8	7
`iiHH`	14	8
`iiiIIIIIII`	23	16
`iiII`	10	8
`iIiI:I`	14	11
`i:--`	11	7
`ii::I(I)`	14	11

4.5 Measuring Metric Complexity

With metric complexity we mean the transition between metric groups, or beats, of two, three, or more pulses. Good examples of a complex metric transition can be found in the bird songs used by Olivier Messiaen, notably in his orchestral work *Chronochromie*.

The metric grouping complexity can be calculated per pattern, phrase, or sentence. It can be applied to any underlying meter as well. One application of this measure is to compare the metric complexity of a particular meter with the grouping complexity of the rhythmic structures that are playing on top of the metric framework. Another application might look at the metric complexity of consecutive rhythmic phrases, as in Messiaen's bird songs, for example.

At the heart of my proposed measurement is a formula that was devloped by Clarence Barlow for his composition system *Autobusk*. In this system, the formula[7] takes into account the rank and the number of primes that take part in the prime number factorization of a natural integer, see Eq. 4.2. Barlow uses this equation in the context of practical tuning systems that need to be quantized to simple ratios.

[7]Barlow calls it the "indigestibility" of a natural number.

Table 4.6 The first 17 values for $\xi(n)$

n	$\xi(n)$
1	0
2	1
3	2.66667
4	2
5	6.4
6	3.66667
7	10.2857
8	3
9	5.33333
10	7.4
11	18.1818
12	4.66667
13	22.1538
14	11.2857
15	9.06667
16	4
17	30.1176

$$\xi(N) = 2\sum_{r=1}^{\infty}\left\{\frac{n_r(p_r-1)^2}{p_r}\right\} \tag{4.2}$$

with $N = \prod_{r=1}^{\infty} p_r^{n_r}$, p_r is the rth prime number and n_r is its exponent in the prime number composition of N. Table 4.6 shows the development of $\xi(n)$ for the first 17 integers.

When sorted after increasing values of $\xi(N)$, the first 17 integers have the following order:

$$\{1, 2, 4, 3, 8, 6, 16, 12, 9, 5, 10, 15, 7, 14, 17\}$$

We will use Barlow's function again in the context of rhythm quantization of musical performances in Chap. 9. For the measurement of metric complexity, I am using the following approach: From a sequence of metric groups, which are measured in number of pulses, the algorithm builds ratios of consecutive groups, reduces them, and calculates, for each reduced ratio $\frac{p}{q}$, the sum $\xi(p) + \xi(q)$. The results are summed together and divided by the number of groups in the sequence. For example, given a sequence of four metric groups, $\{a, b, c, d\}$, and given their reduced ratios $\{\frac{a}{b}, \frac{b}{c}, \frac{c}{d}\}$, $\{\frac{p}{q}, \frac{q}{r}, \frac{r}{s}\}$, then the metric grouping complexity equals:

Table 4.7 The metric grouping complexity of the song thrush that opens *Antistrophe II*

Phrase	Metric complexity	Phrase length
::H	0.333333	8
:'I(I)	0.166667	6
:'I(I)	0.166667	6
:I	0	4
IIIIII(II)	0.2	16
iHiHiH	3.88889	21
i:I	1.22222	7
(HH)	0	8

$$(\xi(p) + 2\xi(q) + 2\xi(r) + \xi(s))/4. \tag{4.3}$$

As an example of the application of the metric grouping complexity, consider Table 4.7. The climax of this bird song, from the rhythmic perspective, is the phrase where the pattern iH gets repeated three times, in pulses: {3, 4, 3, 4, 3, 4}. This phrase has the highest metric grouping complexity, which is immediately followed by a variation of the pattern: i:I, in pulses: {3, 2, 2}. These two phrases stand out because of the 'conflict' between their metric groups, which have a coprime number of pulses. The phrases that only feature groups of 2 or 4 pulses are much smoother in terms of metric complexity.

4.6 The Resolution of Musical Sentences: Effects of Closure and Decline

The resolution of musical sections, including sentences, can be achieved in three ways: 1. through progressive shortening of phrases and patterns, and chunks, which are the basic units that build the rhythmic structure. 2. through progressive prolongation of phrases, patterns and rhythmic chunks. 3. through metrical re-alignment and simplification of rhythmic phrases, which resolve the energy that was previously created in the music, for example by coprime metric groups, or by syncopation.

This example for progressive shortening is taken from Alban Berg's opera *Wozzeck*, 1st Act, bar 76 ff., and it happens in the flutes. After the Hauptmann's phrase (Captain): "so ein Wind macht mir den Effekt wie eine Maus." ("Such a wind always seems to me just like a mouse."), (I):I : :- i H, the flutes answer with .(.): : (.)i'.(.) :'.(.) :(.) '. ($\cdot = \text{♪}$).

The next example is quoted from Béla Bartók, *Musik for String Instruments*, 1st movement, bars 1 - 4. The main theme has the following rhythmic and metric structure ($\cdot = \text{♪}$):

`.X> (.).iX> (.):i:(.) :I:X`,

in pulses: {1, 3, 3|2, 3, 3, 3|3, 3, 3|2, 2, 2, 3}. The last two phrases have both 9 pulses, but the last phrase has 4 groups instead of 3, and it has 8 pitches, one more than the previous phrase. This, together with the descending contour of the melody, creates a closing effect via the second technique of prolongation, but it is also supported by a shortening effect of the metric groups (3 becomes 2).

In Beethoven's *Coriolan* overture, a long coda section at the end of the piece is devoted to a resolution of the entire piece. The energy of the chords recedes through the use of silences, augmentation (longer durations) together with fragmentation, followed by the fading away of the main theme using the same methods. This is the second technique, progressive prolongation together with a reduction of rhythmic density.

Finally, Beethoven's piano sonata no. 13, c-minor, *Pathetique*, 1st movement. The main theme is another good example of how to resolve musical energy (I= ♩):

`IHI IIII IHI IIII HH HH HH HH`

The Greek Amphibrach `.I.` with accentuation of the second note generates a lot of impulse. The resolution starts at the peak of the melody, where the rhythm doubles its values into a chain of half notes. This is the third technique - resolution via realignment and simplification; a chain of longer note values that are aligned with the *alla breve* meter balances the earlier syncopated half of the sentence, which also involved smaller note durations.

4.6.1 Shrinking Durations, or the Accelerando Technique

This is a technique often used by Messiaen. A rhythm is seemingly accelerating by using progressively shorter note durations. Examples are: The descending theme of final piece of *La Nativité du Seigneur*, *Dieu parmi nous*: (· = ♪) `(.)H~I III i W`, or in pulses: $\{(1), 6, 2, 2, 2, 1, 1, 1, 8\}_{24}$. Compare with the three descending trombone themes in *Couleurs de la Cité Céleste* (· = 𝅘𝅥𝅯):

`(I)HH -- . .~H`, $\{(2), 4, 4, 3, 3, 1, 5\}_{22}$

`HH -- .H~I`, $\{4, 4, 3, 3, 1, 6\}_{21}$

`H~.H~.H~. H III .H~I`, $\{5, 5, 5, 4, 2, 2, 2, 1, 6\}_{32}$

It is well known that Messiaen took inspiration for these shrinking rhythms from the Indian deci-talas[8] (Bruhn 2007). In the 13th century treatise, Messiaen found a table with 120 deci-talas. Here are some accelerating patterns: No. 6 `--II.`, no. 53 `HH--` no. 118 `HI.`. A decelerating tala is no. 88 `I-HW`. An analogy from the visual arts is the one-point perspective with structures progressively getting smaller into a single vanishing point.

4.6.2 Triangular Rhythmic Phrases using Primes

Here we look at some examples of rhythmic phrases from the late-romantic period of classical music. The term *triangular* refers to rhythms that have a wave-like triangular structure consisting of a sequence of small, then larger, and again smaller durations, or vice vera. I will compare these rhythms on the basis of their proportions, which is possible after one reduces them to a sequence of integers that represent multiples of the greatest common unit of duration. These proportions make it also easy to transcribe these rhythms into SNMR.

In Peter Illich Tchaikovsky's, symphony no. 6, 1st movement, the development section brings the following *ostinato* figure in the horns (at rehearsal letter M):

`(I)II~--~III~--~III~--~I`

It translates to the following durations - based on the greatest common pulsation: $\{(2), 2, 5, 5, 2, 5, 5, 2, 5, 5\}_{38}$. Later on (3 bars before letter Q), we find the variation: `III~--` with the durations $\{2, 2, 5, 3\}_{12}$. Compare with Richard Strauss, *Tod und Verklärung* (*Death and Transfiguration*). The beginning in the timpani introduces the following rhythmic theme: `(-)-~III~--~III~--~I` It has the durations $\{(3), 5, 2, 5, 5, 2, 5, 5\}_{32}$, and it seems not far away from of the Tachikovsky rhythm, only much slower in tempo. In Richard Wagner's *Tristan und Isolde*, act 2, the scene "O sink hernieder, Nacht der Liebe", has the following rhythm in the accompaniment: `IHHI~--`, duration $\{2, 4, 4, 5, 3\}_{18}$. The Ring has Alberich characterized with `(.)-.III(.)`, or in durations: $\{(1), 3, 1, 2, 2, 2, (1)\}_{12}$. The prolog of Götterdämmerung, act 2, varies this rhythm of unrest in many ways. Anton Bruckner, symphony no. 8, 3rd movement, starts with `IH~--IH~--`, which contains only primes as durations: $\{2, 7, 3, 2, 7, 3\}_{24}$, and the well-known Bruckner rhythm `--III` uses the smallest primes $\{3, 3, 2, 2, 2\}_{12}$. It is interesting that these late-romantic rhythms employ mainly prime numbers. The list can go on with Alban Berg in his violin concerto: `-.I--`, or $\{3, 1, 2, 3, 3\}_{12}$ or with Gustav Mahler's beginning of his symphony no. 9, 1st movement, `-.(.)-`, in duration $\{3, 1, (1), 3\}_8$, or simplified as `-I-`, which is equal to $\{3, 2, 3\}_8$.

[8] Śārṅgadeva's Saṅgita Ratnākara.

4.7 The Sentence Algorithm in *Chunking*

The *sentence* module in *chunking* generates the rhythmic structure of musical sentences, as a blue-print for further compositional work. It implements the asymmetry of the phrases and patterns that we have analyzed so far. It uses the principles of coprimality of neighbouring structures, and it follows the triangle principle we have discussed in the context of organic form.

As with all generative algorithms there are certain limitations, for example the maximum total length of the sentence is 120 pulses. The limits haven been found after experimentation. If the user wishes to extend them, then it would only require some changes in the source code to do so. Secondly, because the method uses integer partitions into k distinct parts, the maximum allowed number of parts per structural unit is 5. The output is a series of recursive partitions of n into k parts. For the resulting parts, I am using the following hierarchy of all musical components: *Sentence*, *Phrases*, *Patterns* and *Chunks*. After one has decided on the final length of the sentence n, this integer is partitioned into k phrases, which, in turn, are further partitioned into at most k patterns. The patterns, finally, are partitioned into chunks of 2s and 3s. The partitions of all of these components are always picking coprime lengths. Here is an example for $n = 32$ and $k = 3$:

$$Sentence\{32\},$$

$$Phrases\{9, 13, 10\}_{32},$$

$$Patterns\{\{5, 4\}_9, \{6, 7\}_{13}, \{3, 5, 2\}_{10}\}_{32},$$

$$Chunks\{\{\{3, 2\}_5, \{2, 2\}_4\}_9, \{\{2, 2, 2\}_6, \{3, 2, 2\}_7\}_{13}, \{\{3\}_3, \{3, 2\}_5, \{2\}_2\}_{10}\}_{32}$$

It was found, after experimentation, that the coprimality between individual parts of a rhythmic sentence is a condition for good musical results from partitions, because of the balance of impulse and resolution that is generated. Note also that the partitions of 32 and of 10 have their distinct parts arranged as a musical triangle, i.e. the highest figure has been placed by the algorithm into the center of the sequence. Figure 4.1 shows a possible transcription based on the above structure of the 32-pulse sentence. In SNMR it would read:

```
w I v : v : v i I I > - I I
```

In order to generate balanced musical triangles, the algorithm searches for the partition where difference between its parts is relatively small. This is achieved by making sure that the standard deviation, σ, is as small as possible. For example, if $n = 24$ and $k = 3$ one finds that the partition $\{7, 8, 9\}_{24}$ meets all conditions: Coprimality of 3 distinct parts, and the parts are closest together, i.e. σ converges

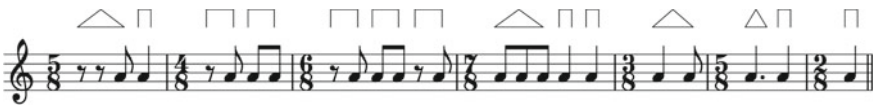



Fig. 4.1 One of many possible transcriptions that are generated by the *sentence* algorithm. Above the staff are conductor signs. The triangle indicates a ternary group, and the open square means a binary group

to a minimum in such cases. The *sentence* algorithm has a built-in preference for choosing a partition like this for the generation of phrase and pattern lengths.

A partition alone does not determine any order among the parts. This is different for music where the order of events does matter, because it is a result of the continuous transmission of moments, i.e. the chain of causes and effects, which lead to a musical experience (see Chap. 2). For that reason, the algorithm reorders the parts of a partition into a musical triangle. It means that a musical triangle ranks the order of the size of values in such a way that the largest one is near or at a central position, which marks the maximum expansion. The size of the values are the durations of the different parts of a sentence that are measured in pulses. Around the central peak, the smaller values are grouped in various ways, either leading up towards the summit, or descending thereafter. For example, {1, 2, 4, 3}, or {2, 3, 4, 6, 5, 1}. The *sentence* algorithm knows fourteen shapes of this kind for the triangular ordering of all partitions from three to seven distinct parts. Table 4.8 shows the complete list, where L denotes the longest element that should be at or near the center, and where s_1 and s_2 denote the shortest and the second-shortest element in the set; w, x, y, z are the potential intermediate values of durations in lexicographic order. Partitions with only two distinct parts have only two obvious solutions.

After the ordering of the phrases according to the list of musical triangles, the next lower structural level is created from patterns, which are also ordered according to one of the triangle shapes. The selection of these shapes follows the number of parts, otherwise it is random.

The lowest level of the process is the creation of chunks from each of the patterns. A chunk can only last for two or three pulses, therefore at this level, the algorithm operates on partitions of n into 2s and 3s. For example, if a pattern has the length 8, then the only partition into 2s and 3s is $\{3, 3, 2\}_8$. When n is larger, many different kinds of these partitions can emerge. For $n = 70$ one already counts 11 partitions into 2s and 3s. These partitions will naturally have many repetitions of 2s and 3s, which distinguishes them from the partitions into k distinct parts. I call the partitions into 2s and 3s *templates*, because of the fair number of possible orderings they can bring about. Templates offer interesting combinations. For example, $\{3, 3, 3, 2\}_{11}$ can be rotated three times: $\{3, 3, 2, 3\}_{11}$, $\{3, 2, 3, 3\}_{11}$, and $\{2, 3, 3, 3\}_{11}$. When played as a sequence, these rotations form a palindrome, i.e. the same sequence emerges when played backwards. Each rotation on its own though creates a different metric feel.

A template, i.e. a partition of n into 2s and 3s, when rotated into its lowest lexicographic position, generates a special mathematical structure called *bracelet* (Karim

Table 4.8 Fourteen types of musical triangles for partitions with k distinct parts. The index lists assume that the original partitions are sorted from the smallest to the highest value

k	Triangular order	Indices
3	$\{s_1, L, s_2\}$	{0, 2, 1}
	$\{s_2, L, s_1\}$	{1, 2, 0}
4	$\{s_1, x, L, s_2\}$	{0, 2, 3, 1}
	$\{s_1, L, x, s_2\}$	{0, 3, 2, 1}
	$\{s_2, x, L, s_1\}$	{1, 2, 3, 0}
	$\{s_2, L, x, s_1\}$	{1, 3, 2, 0}
5	$\{s_1, x, L, y, s_2\}$	{0, 2, 4, 3, 1}
	$\{s_1, y, L, x, s_2\}$	{0, 3, 4, 2, 1}
	$\{s_2, x, L, y, s_1\}$	{1, 2, 4, 3, 0}
	$\{s_2, y, L, x, s_1\}$	{1, 3, 4, 2, 0}
6	$\{s_1, x, z, L, y, s_2\}$	{0, 2, 3, 5, 4, 1}
	$\{s_2, x, z, L, y, s_1\}$	{1, 2, 3, 5, 4, 0}
7	$\{s_1, y, z, L, x, w, s_2\}$	{0, 4, 5, 6, 3, 2, 1}
	$\{s_2, y, z, L, x, w, s_2\}$	{1, 4, 5, 6, 3, 2, 0}

Table 4.9 All bracelets for $n = 16$ containing only the integers 2 and 3. They are based on the two partitions of 16 that have only 2 and 3s as parts, i.e. two templates of 16. Each bracelet can be rotated and reversed, thereby creating all possible combinations of the parts within a template

{2, 3, 3, 2, 3, 3}
{2, 3, 2, 3, 3, 3}
{2, 2, 3, 3, 3, 3}
{2, 2, 2, 3, 2, 2, 3}
{2, 2, 2, 2, 3, 2, 3}
{2, 2, 2, 2, 2, 3, 3}

et al. 2013). A bracelet is a cyclic pattern with a specific sum of its parts, but it is *unique* with respect to its rotations and reversals. No other bracelet with the same sum of its parts would emerge under all the rotations and reverse orders. When we write a template as a bracelet, then $n = 11$ results in the following two forms: $\{2, 3, 3, 3\}_{11}$ and $\{2, 2, 2, 2, 3\}_{11}$. However, there are templates, such as $\{2, 2, 2, 3, 3\}_{12}$, that spawn more then one bracelet; the other form is $\{2, 2, 3, 2, 3\}_{12}$. Table 4.9 shows all bracelets for $n = 16$. Three bracelets[9] each are possible for both templates of 16.

To summarize, the *sentence* algorithm uses two kinds of partitions: 1. Partitions of n into k distinct parts for organizing a sentence into phrases and patterns. 2. Partitions of n into sequences of 2s and 3s, which control the use of rhythmic chunks within patterns. These are the metric groups of 2s and 3s, which I call templates. When sorted into lexicographical order, templates are equivalent to bracelets. However,

[9]Karim et al. (2013) published an efficient algorithm for generating bracelets.

there can be more than one form of bracelet per individual template. Bracelets have the nice property that they are unique in terms of the sum of their elements, and in terms of their rotations and reverse orderings. A template can, therefore, spawn a large number of rhythms that are based on a specific pattern, i.e. a metric grouping of our chunks that 2 or 3 pulses in length. These possibilities lead to the following classification of patterns according to their musical effect.

4.7.1 Seven Categories of Rhythmic Patterns

So far, I have found seven categories for all the different rhythmic patterns that are based on metric chunks of 2 and 3 pulses. I have named them *resistor, release, arch, catenary, growth, decline,* and *alternating*. All patterns have been generated by taking a bracelet as a starting point, and by working on it with the operations of rotation and reversal. The names of the categories reflect the musical forces of impulse and resolution that are present within these patterns. I have mentioned already that the order of events can never be arbitrary because of the inherent cause-and-effect relationship between musical events.[10]

The *resistor* pattern consists of one or more 2s followed by one or more 3s. For example, Table 4.9 contains the bracelet $\{2, 2, 3, 3, 3, 3\}_{16}$. This category was inspired by the phenomenon that the transition from a pulsation of 2s into a pulsation of 3s creates a resistance against the flow of the binary pulsation. One could also think in analogy of the electronic resistor, in which resistance is built against the flow of electric charge. Similarly, the transition from a binary to a ternary pulsation builds up musical impulse, which might resolve later on by switching back to a binary pulsation, for example, when the pattern is repeated. In this repeated manner, musical impulse and resolution can interchange like the peaks and valleys of a wave.

The *release* category is the opposite form of the *resistor* pattern: One or more 3s are followed by one or more 2s, for example $\{3, 3, 2, 2, 2, 2, 2\}_{16}$; see also the last bracelet in Table 4.9 in reverse order.

The combination of a *resistor* pattern with a *release* pattern creates an *arch* pattern. I call this category *arch*, because the bigger number 3 is surrounded by sequences of one or more 2s. In this category, impulse and release are combined within one and the same pattern, similar to a pendulum being pushed out of equilibrium and then returning backwards. Consider, for example, $\{2, 2, 3, 3, 2, 2, 2\}_{16}$ which is a rotation of the last bracelet in Table 4.9.

If one combines an *arch* pattern with a *release* pattern, an upside-down arch is created. I call this category, the *catenary* pattern. A catenary is the curve produced

[10]Causality is also found in the physicality of musical events. For example, the action of plucking a string generates a series of harmonics. It is also inherent to our perception of musical events. For example, a melody is not perceived as a string of single unrelated sounds, but as a series of tones that participate *together* to form a coherent musical line. This perception of coherence triggers a sense of unity, and it is dependent on the causes that arose from the physical parameters (pitch, loudness, tempo, and intonation) that have been controlled by the musician during the performance.

by a hanging chain or cable that is only supported at its ends. In our case, the integer 2, or a series of 2s, is surrounded by 3s, as for example in $\{3, 2, 2, 3, 3, 3\}_{16}$, which is a rotation of the third bracelet in Table 4.9. The number of elements in each category can vary, but in *chunking* it can only go up to 9. As a consequence, the total lengths in pulses within a category also varies, but it is limited to 26 pulses.

Growth and *decline* also form opposite categories. In a *growth* pattern, one of the elements, either 2 or 3, is allowed to grow in a sequence: Consider, for example, $\{2, 3, 2, 3, 3, 3\}_{16}$, the second bracelet in Table 4.9. In it, the smallest resistor pattern, $\{2, 3\}_5$, grows into $\{2, 3, 3, 3\}_{11}$, via a repetition of 3s. To continue the growth, the next instalment could bring $\{2, 3, 3, 3, 3\}_{14}$. In an obvious variation of the same principle, one could also have the case where a series of 2s grows gradually from $\{2, 2, 3\}_7$ to $\{2, 2, 2, 3\}_9$, which is a rotation of the fourth bracelet in Table 4.9. Or, both elements can develop at the same time, from $\{2, 3\}_5$, over $\{2, 2, 3, 3\}_{10}$, to $\{2, 2, 2, 3, 3, 3\}_{15}$. In all cases, the *growth* pattern can be generated by sequencing a selection of, either *resistor* patterns getting longer and longer, or by taking *release* patterns of increasing lengths.

As one might expect, the *decline* pattern behaves like the *growth* pattern, only in the opposite direction. One can construct *decline* patterns by sequencing either *resistor* patterns, or *release* patterns whilst observing that their lengths gradually decrease.

In the implementation of the *sentence* algorithm, the maximum length for a *growth* pattern within a phrase is twenty-five pulses with a maximum number of nine elements. This is another consequence of applying Miller's 'magic number' 7 ± 2 within the context of musical sentences. Outside of the *sentence* context, one could compose a section of a piece by sequencing patterns taken directly from any of the seven categories. This is easily possible, of course, and one would still observe Miller's 7 ± 2 on a local level, because all patterns have never more than 9 elements.

The final category is the collection of *alternating* patterns. The first bracelet in Table 4.9 is such a pattern. In general, an *alternating* pattern is simply a repetition of a smaller pattern, but with the aforementioned restriction that there are no more than 9 elements in a pattern like this.

As an extension to Arom's Aksak theory, we can confirm that the patterns *resistor, release, arch* and *catenary* appear in the list of all Aksak rhythms, and they appear also in sequences. Aksak rhythms with a $length > 20$ pulses can be conceived as complete rhythmic sentences. The calculation of bracelets with fixed content over the alphabet $\{2, 3\}$ will generate all patterns, from which rotations and reversal operations can generate all possible Aksak rhythms.

4.7.2 Transcription of Patterns and the Complete Sentence

The *sentence* algorithm generates a series of outputs in the form of SNMR, *lilypond* scripts, and *Csound* scores. The patterns of the seven categories are broken down into actual rhythmic chunks by using the SNMR's possibility to match any binary

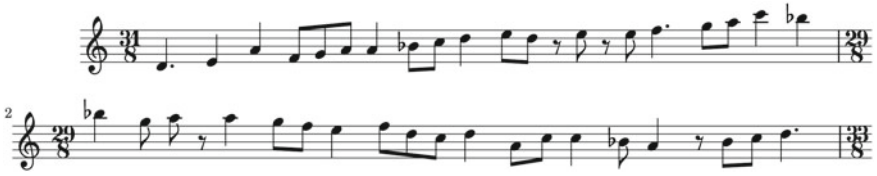



Fig. 4.2 A sentence with 60 pulses generated with *chunking*. Its internal structure is shown below in Table 4.10

Table 4.10 Example structure of a sentence

$Sentence\{60\}$
$Phrases\{31, 29\}_{60}$
$Patterns\{\{16, 15\}_{31}, \{14, 15\}_{29}\}_{60}$
$Chunks\{\{\{3, 2, 2, 3, 2, 2, 2\}_{16}, \{3, 3, 3, 2, 2, 2\}_{15}\}_{31}, \{\{3, 2, 2, 2, 2, 3\}_{14}, \{2, 2, 3, 3, 2, 3\}_{15}\}_{29}\}_{60}$

or ternary chunk to a metric group of either 2 or 3 pulses. Remember the list of binary and ternary rhythm chunks that was shown in Table 3.4. To pick from any of the binary or ternary chunks, *sentence* uses a weighted random choice. Although the actual weights are a matter of personal preference, my experiments have shown that the probability of chunks that start with a rest had to be reduced in favor of those chunks that start with an onset. I found that if there would be an equal chance, then it would lead to more isolated chunks and it would break up the rhythmic flow of the music. Subsequent chunks would not connect well enough with each other in order to form the next higher level from a unity of patterns, to a unity of phrases, reaching up to a unity of the whole sentence. A weighted probability for each chunk seemed to be necessary to achieve a better solution.

I gave a practical example for the transcription into a *lilypond* script by showing the sentence in Fig. 4.1. The above Fig. 4.2 shows you yet another good example; here, I added melodic lines to the rhythmic structure that was produced with the *sentence* module. Table 4.10 demonstrates its structure:

4.8 Conclusion

Chunking's *sentence* algorithm creates rhythmic sentences that subdivide into phrases, patterns of binary and ternary chunks, and it generates the specific composition of rhythms based on these chunks. The algorithm uses a top-down approach based on partitions, bracelets and seven musical categories of patterns. It adheres to Miller's magic number on multiple levels and it uses the concept of chunks and templates for musical variation. The output in a general notation format helps to navigate choices during the composition process. That process does not stop with the output of *chunking*, but uses its material to compose entire pieces of music.

I also have analyzed sections of music by Stravinsky and Messiaen. It was found that the strategy of coprimality between metric groups was used as an important principle of form. In addition I introduced an algorithm for the measurement of metric complexity, which gives information about the level of impulse that is generated through the use of asymmetry in rhythmic sentences.

References

Arom, S. (1991). *African Polyphony and Polyrhythm*. Cambridge: Cambridge University Press. (ISBN 052124160X).

Baraldi, F. B., Bigand, E., & Pozzo, T. (2015). *Empirical Musicology Review*, *10*(4), 265–291. https://doi.org/10.18061/emr.v10i4.4891.

Bouroubi, S., & Tani, N. B. (2009). Integer partitions into arithmetic progressions with an odd common difference. *Integers*, *9*, https://doi.org/10.1515/INTEG.2009.008.

Brăiloiu, C. (1951). Le Rythme Aksak. *Revue de Musicologie*, *33*(99 and 100), 71–108. http://www.jstor.org/stable/926002.

Bruhn, S. (2007). *Messiaen's Contemplations of Covenant and Incarnation: Musical Symbols of Faith in the Two Great Piano Cycles of the 1940s*. Hillsdale, NY: Pendragon Press.

Celibidache, S. (2008). *Über Musikalische Phänomenologie*. Augsburg: Wissner-Verlag. (ISBN 978-3-89639-641-9).

Fischer, M., Holland, D., & Rzehulka, B. (1986a). Gehörgänge. Zur Ästhetik der musikalischen Aufführung und ihrer technischen Reproduktion. In (chap. Musik verschwindet. Gespräch der Autoren mit Sergiu Celibidache 1985). Peter Kirchheim. (ISBN 3874100162).

Fischer, M., Holland, D., & Rzehulka, B. (1986b). Gehörgänge. zur ästhetik der musikali-schen aufführung und ihrer technischen reproduktion. In (chap. Sergiu Celibidache im Gespräch mit Joachim Matzner). Peter Kirchheim. (ISBN 3874100162).

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*(1), 1–40. http://www.sciencedirect.com/science/article/pii/S0010028596900110, https://doi.org/10.1006/cogp.1996.0011.

Hardy, G., & Wright, E. (2008). *An Introduction to the Theory of Numbers* (6th ed.). Oxford: Oxford University Press.

Junius, M. (1983). *Die Tālas der Nordindischen Musik*. Unterwössen, Bavaria, Germany: Katzbichler, Bernd. (ISBN 978-3873973046).

Karim, S., Sawada, J., Alamgir, Z., & Husnine, S. (2013). Generating bracelets with fixed content. *Theoretical Computer Science*, *475*, 103–112.

Knuth, D. E. (2013). *The Art of Programming. Combinatorial Algorithms. Part 1* (4th ed.). New Jersey: Addison-Wesley. (ISBN 0-201-03804-8).

London, J. (2012). *Hearing in Time. Psychological Aspects of Musical Meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Messiaen, O. (1963). *Chronochromie pour Grand Orchestre*. Paris: Alphonse Leduc.

Miller, G. H. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Polak, R. (2010). Rhythmic feel as meter. Non-isochronous beat subdivision in jembe music from Mali. Music Theory. *Online*, *16*(4),

Polak, R. (2015). Pattern and variation in the timing of aksak meter: Commentary on Goldberg. *Empirical Musicology Review*, *10*(4), 329–340. https://doi.org/10.18061/emr.v10i4.4883.

Reina, R. (2014). *Karnatic rhythmical structures as a source for new thinking in western music* (Doctoral dissertation, Brunel University, London, UK). http://www.rafaelreina.org/doctorate.html. Retrieved March 2018.

Sankaran, T. (2010). *The Art of Konnakkol (Solkattu)*. Toronto, Canada: Lalith Publishers. (ISBN 978-0-9698823-1-2).

Thakar, M. (2011). *Looking for the Harp Quartet: An Investigation into Musical Beauty*. Rochester, NY: University of Rochester Press. (ISBN 978-1-58046-346-1).

Wilf, H. S. (2000). *Lectures on Integer Partitions* (Technical Report). Philadelphia: University of Pennsylvania. https://www.math.upenn.edu/~wilf/PIMS/PIMSLectures.pdf. Retrieved March 2018.

Chapter 5
The Use of the Burrows–Wheeler Transform for Analysis and Composition

5.1 Introduction

In this chapter, I introduce a new device for analyzing musical rhythms and for music composition in general: The Burrows–Wheeler Transform (BWT). It has many applications in the area of bioinformatics. It is used in the research of genomes, for the detection of sequences within the genome, and for the detection of mutations and genetical diseases. In bioinformatics, a genome is represented as a string of the nucleobases that build the structure of DNA. For example, the string ACGT represents the four bases found in DNA: Adenine, cytosine, guanine, and thymine. One can regard the genome as one very long word over the alphabet $\{A, C, G, T\}$. In the case of the human genome, it is a word with over 3 billion letters. Because the complete genome information uses a large amount of computer memory, efficient algorithms, such as the BWT, have been developed in order to keep the genome data compressed, whilst, at the same time, performing exact and inexact pattern-matching. It is a very efficient method when searching for mutations, for example.

The BWT is a not a compression algorithm. It is used, however, to obtain better compression rates, because strings transformed with BW are better candidates for run-length encoding (RLE). After applying the BWT to a word, the letters of a string have the tendency to cluster together. Because of cyclic rotations and lexicographic sorting, the BWT contains valuable information about all substrings, their positions and their rotations in a highly compact form that is equal to the length of the string (Fenwick 2007). A suffix of a word, such as a genome, is created by taking characters away from the beginning of the word. The remainder of the word is the suffix. If the word has length n, then there are n suffixes in an array. A suffix array might need a significant amount of memory, as Polanski and Kimmel (2007) point out:

> "The suffix array for the human genome constitutes approximately 12 gigabytes (3 billion 4-byte integers) of RAM. However, the BW string alone is sufficient to determine word counts and can be compressed to about 1 gigabyte of RAM. Furthermore, for the purpose of querying, all but a negligibly small portion of the compressed form can remain so throughout execution. Using the BW transformation, any region of the genome can be annotated with its constituent k-mer frequencies."

Table 5.1 The BWT matrix of the word *Saskatchewan*. The last column (read from top to bottom) is the BWT word

1	`Saskatchewan`
2	`anSaskatchew`
3	`askatchewanS`
4	`atchewanSask`
5	`chewanSaskat`
6	`ewanSaskatch`
7	`hewanSaskatc`
8	`katchewanSas`
9	`nSaskatchewa`
10	`skatchewanSa`
11	`tchewanSaska`
12	`wanSaskatche`

A k-mer is a substring of length k. This technique has similar applications within the context of musical rhythms. A rhythm can be regarded as a kind of genome that is built out of simple constituent elements, an alphabet of chunks. We have previously discussed the use of chunks in Chaps. 3 and 4. With regard to rhythm, our interest is in finding and isolating patterns, i.e. k-mers, within a larger sequence.

5.2 The BWT Algorithm

In this section, we investigate the use of the BWT for rhythm analysis. Imagine that the durations of a rhythmic pattern is equivalent to a word over a certain alphabet. Our SNMR is a good example: One ASCII symbol per rhythmic chunk.

The algorithm of the Burrows–Wheeler transform works as follows: First, a word w, which is a string of characters from a defined alphabet, will be rotated character for character from back to front, i.e the last character becomes the first, then the new string's last character becomes the first, and so forth. There are $|w| - 1$ possible rotations, which are called the *non-trivial conjugates* of the word, with the length of the word notated as $|w|$.

The BWT of w arranges all of its non-trivial conjugates into a matrix in lexicographic order (Berstel 2009). The last column of the matrix is the word resulting from the transform: $BWT(w)$. See the BWT matrix of the word *Saskatchewan*[1] in Table 5.1. Its BWT word is the last column: **nwSkthcsaaae**.

Because of its tendency to cluster the letters of strings together, the Burrow-Wheeler transform is used in the open-source lossless data compression algorithm bzip2[2] (Burrows and Wheeler 1994). The inverse process of the BWT rebuilds the

[1] Saskatchewan is a river in Canada and the name of one of its provinces.

[2] see http://bzip.org.

Table 5.2 The reconstruction of the word *Saskatchewan* by creating the inverse BWT matrix. The starting point is the first column on the left, which is the BWT word **nwSkthcsaaae** sorted lexicographically. Table 5.3 shows the final matrix in column 12

1	2	3	4	5	6	7	8	9	10
S	Sa	Sas	Sask	Saska	Saskat	Saskatc	Saskatch	Saskatche	Saskatchew
a	an	anS	anSa	anSas	anSask	anSaska	anSaskat	anSaskatc	anSaskatch
a	as	ask	aska	askat	askatc	askatch	askatche	askatchew	askatchewa
a	at	atc	atch	atche	atchew	atchewa	atchewan	atchewanS	atchewanSa
c	ch	che	chew	chewa	chewan	chewanS	chewanSa	chewanSas	chewanSask
e	ew	ewa	ewan	ewanS	ewanSa	ewanSas	ewanSask	ewanSaska	ewanSaskat
h	he	hew	hewa	hewan	hewanS	hewanSa	hewanSas	hewanSask	hewanSaska
k	ka	kat	katc	katch	katche	katchew	katchewa	katchewan	katchewanS
n	nS	nSa	nSas	nSask	nSaska	nSaskat	nSaskatc	nSaskatch	nSaskatche
s	sk	ska	skat	skatc	skatch	skatche	skatchew	skatchewa	skatchewan
t	tc	tch	tche	tchew	tchewa	tchewan	tchewanS	tchewanSa	tchewanSas
w	wa	wan	wanS	wanSa	wanSas	wanSask	wanSaska	wanSaskat	wanSaskatc

Table 5.3 The reconstruction of the word *Saskatchewan* continued. The highlighted substrings in column 12 build the important suffix array. The suffix array indices on the right are used to find the positions of matching substrings in the word

11	12	Suffix#
Saskatchewa	**Saskatchewan**	0
anSaskatche	**an**Saskatchew	2
askatchewan	**askatchewan**S	11
atchewanSas	**atchewan**Sask	8
chewanSaska	**chewan**Saskat	6
ewanSaskatc	**ewan**Saskatch	4
hewanSaskat	**hewan**Saskatc	5
katchewanSa	**katchewan**Sas	9
nSaskatchew	**n**Saskatchewa	1
skatchewanS	**skatchewan**Sa	10
tchewanSask	**tchewan**Saska	7
wanSaskatch	**wan**Saskatche	3

original string in the following way: The BWT word is written down as a column and will be sorted lexicographically. This new column is the first column of the original matrix. Then, the BWT word is prepended as the new first column. The rows of this new matrix are again sorted lexicographically. Because of this sort, the new first column becomes again the the first column of the initial BWT matrix. These steps are repeated until a matrix of size $|w|^2$ has been created where the rows are in lexicographic order. In addition, the suffix array is also sorted lexicographically in the BWT matrix, see table. See Tables 5.2 and 5.3 for an example.

It is characteristic for the iBWT to visit all factors, i.e. substrings, of the circular word in lex order. This is most interesting from the musical point of view because the iBWT process creates a systematic overview of the cyclic combinatorial changes

for any given rhythmic pattern, which helps to gain a deeper knowledge of the compositional material. Progressive shortening or expansion of patterns are key concepts in classical music theory (Caplin 2000). The recombination of substrings can model processes of generating new words out of existing ones. This would include modelling growth, stasis and decline of the material, which are important tools for establishing the form of a musical piece, as well as the form of all of its constituent elements. See Chap. 4 for a demonstration of this technique.

As a tool for rhythm analysis, the BWT allows us to find relationships between patterns of different lengths, including also patterns that have been rotated. With the help of the iBWT process one can obtain the building blocks for organic growth and decline of musical themes and motives. The separation of motives from longer units, like phrases, can be modelled here. The iBWT allows for a deconstruction of the original musical pattern. By selecting material from the iBWT matrix one can establish a relation of unity between smaller patterns and larger scaled musical entities, such as musical sentences, for example. For the composer, the iBWT process can serve as a tool to establish coherence in musical compositions.

5.2.1 The Inverse BWT Algorithm (iBWT)

The iBWT process builds all permutations of substrings of the input that are also present as a substring in the *cyclic* word. The order of the lists of substrings is lexicographic, with only two elements changing position at a time, when advancing from one substring to the next, similar to a Gray code. Depending on the amount of repetitions in the original string, there will be redundancy in the lists of substrings. For compositional purposes, as the composer might want to have only the strings that are unique, the methods used in our *chunking* application[3] are able to filter out those repetitions. The lexicographic order of the iBWT matrices is also useful because the beginning of neighbouring rows will be similar but also different. For composition this is useful, because it will allow the composer to play with the information of a motive or a phrase that is well remembered, for example.

There is an even more efficient algorithm for the iBWT, which uses the so-called *first-last property*: The ith occurrence of character c in last column is the same text character as the ith occurrence of c in the first column. Using this property it is possible to re-construct the original word by just using the BWT word and the first column of the BWT matrix. The first column of the matrix is the BWT word sorted lexicographically.

[3] https://github.com/gboenn/chunking.

5.2.2 A Rhythm Analysis Program Using the BWT

In *chunking*, the *anaphrases* module implements the analysis of a rhythmic pattern by using the BWT.

1. Rhythmic sequences in SNMR are translated into words over the alphabet $A = \{a, b\}$, in such a way that note onsets are encoded with the letter a, whilst all other pulses that happen in-between are encoded with the letter b. Therefore, a sequence of bs represents inter-onset *pulses*. For example, the rhythm `-IX>` translates into the word $\{a, b, b, a, b, a, a, b, a, b, a\}$. We call the latter a *rhythm word*.
2. The algorithm then performs the BWT for each *rhythm word*.
3. Using the BWT word one can re-construct the original *rhythm word* with very simple sort and indexing procedures. In this reconstructing process, all cyclic substrings of the original *rhythm word* [4] can be visited and matched against any query. If we call the length of the word w that enters the BWT $|w|$, then this process results in a series of $|w|$ matrices containing every *k-mer*, with $0 < k \leq |w|$. In the context of rhythm, the substrings are not formed at the original position within the metric grid of pulses beneath the original rhythm, but they appear as metric shifts in relation to the original rhythmic sequence. That means the metric feel of the original rhythm changes if the elements of the rhythm are rotated, and if cyclic sequences are cut out of this rhythm. Examples for this process are given in the next section.

5.2.3 Fragmentation Modelling by Using the iBWT Matrix

In terms of organic form, a decay of the rhythmic material can be composed with the iBWT matrix, starting with long substrings (k-mers), getting increasingly shorter until the original rhythm is dissolved. This process is also called fragmentation in music analysis, or dissociation.

Another aspect of organic form is the new growth that comes out of decay. The new combinations can form directly from the material of the iBWT matrix. Using the example of Table 5.4, one can generate new rhythmic phrases by combining row 13 with row 8: `XIIX> IXX` . Another possibility for new growth is the rotation of the phrase, for example using row 11, see Table 5.5.

The development of an initial rhythmic sentence can be composed by using fragmentation and repeption, see Table 5.6.

Another form of development can be achieved by using fragmentation, rotation and repetition together, see Table 5.7. Rows 3, 4 and 5 are sequential rotations of the short fragment `IIX:` in row 2.

[4] In bioinformatics one calls the substrings *k-mers*.

Table 5.4 Example of fragmentation of a 36 pulse long rhythmic theme. The lengths of the fragments are shown in pulses on the left

36	IX IIX IIIXX IIX IX
22	IXX IIX IXI
13	XII X>
8	IXX
5	I>
3	>
2	v
1	.

Table 5.5 Example of rotation as a means to develop material.

0	I> >+
1	> >+I
2	II:I>
3	I:I>I
4	:I>II

Table 5.6 Example of fragmentation and repetition used with shorter substrings of the original rhythmic sentence

0	36	IX IIX IIIXX IIX IX
1	2 × 19	<XIXIX> <XIXIX>
2	4 × 9	IIX: IIX: IIX: IIX:
3	3 × 5	< : < : < :
4	5 × 3	> > > > >
0	36	IX IIX IIIXX IIX IX

Table 5.7 Fragmentation, rotation and repetition based on a single rhythmic sentence, thereby generating a section of musical development

0	36	IX IIX IIIXX IIX IX
1	2 × 19	<XIXIX> <XIXIX>
2	2 × 9	IIX: IIX:
3	2 × 9	IX:I IX:I
4	2 × 9	X:II X:II
5	2 × 9	:IIX :IIX
6	3 × 8	<:> <:> <:>
0	36	IX IIX IIIXX IIX IX

5.3 Conclusion

The iBWT matrix is an ideal algorithm for building shorter rhythm phrases out of a longer sentence. Thus it can be used to model fragmentation, i.e. decay in organic form. Growth is generated from 'decayed' material through the processes of recombination, repetition and rotation. Through the application of the techniques above, one has the means to compose and to develop sections out of previously presented rhythmic or thematic material. In order to find shorter sub-sequences

(k-mers) in iBWT matrices, one can use heuristics, experimentation, proportional series (golden section), integer partitions (as discussed in the previous Chap. 4) and other kinds of arithmetic approaches. The following chapter uses the BWT for the analysis of a collection of rhythmic patterns that originate from a very special type of rhythm word, the so-called *Christoffel words*.

References

Berstel, J. (2009). *Combinatorics on words: Christoffel words and repetitions in words* (Vol. 27). CRM monograph series. Providence: American Mathematical Society. (ISBN 9780821873250).

Burrows, M., & Wheeler, D. J. (1994). *A block-sorting lossless data compression algorithm* (Technical Report). Palo Alto: Digital Systems Research Center. (SRC Research Report 124).

Caplin, W. E. (2000). *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven*. Oxford: Oxford University Press. (ISBN 019514399X).

Fenwick, P. (2007). Burrows–Wheeler compression: Principles and reflections. *Theoretical Computer Science*, *387*(3), 200–219. https://doi.org/10.1016/j.tcs.2007.07.012.

Polanski, A., & Kimmel, M. (2007). *Bioinformatics*. Berlin: Springer. (ISBN 978-3-540-24166-9).

Chapter 6
Christoffel Rhythms

6.1 Introduction

Christoffel words[1] have been known to the field of mathematics since Jean Bernoulli in the late 18th century. Today, they have gained specific interest in the area of combinatorics on words (Berstel 2009; Glen 2012; Lothaire 2005). This chapter will show that they are also useful for music analysis and composition. Remember that a word is an ordered, finite or infinite sequence of symbols taken from a finite alphabet that is the non-empty set A. Transcribing a Christoffel word into a rhythmic pattern is very simple. It only involves a mapping of the letters of the word to note onsets and silences. An analysis of the resulting rhythmic patterns indicates that they are strongly related to archetypical classical rhythms and to world music rhythms. When working with Christoffel words, one discovers similarities to African and Latin-American music, Minimalism, and other musical styles that feature non-symmetric meters and rhythms. The links with rhythms and meters of many different cultures provide a fertile ground for musical exploration and discovery. We also show that there exists a unique Christoffel word, or one of its conjugates, for every Euclidean rhythm (Demaine et al. 2009). We will present a structural comparison between Christoffel words and Euclidean rhythms.

Here are a few conventions followed in this chapter. Christoffel words are constructed over the two-word alphabet $\{a, b\}$, with $a \prec b$. For the purpose of rhythm notation, resulting words are being mapped with $a \mapsto 1, b \mapsto 0$. We call the result of this mapping a Christoffel rhythm. The 1s represent note onsets, 0s represent silent pulses that fall in the inter-onset-times. Although common practice, this mapping is not unique; therefore, we will also present several bit-operations on rhythmic patterns, which produce characteristic counterpoints to the standard mapping.

[1] Named after the German mathematician Elwin B. Christoffel (1829–1900).

6.2 Christoffel Rhythms from Christoffel Words

A Christoffel word $C(m, n)$ is defined by a pair of coprime integers, $m \perp n$. For example, if the word is constructed over the two-letter alphabet $\{a, b\}$, then $C(3, 4) = aababab$. After mapping, this word expresses a seven pulse rhythmic pattern: $\{1, 1, 0, 1, 0, 1, 0\}$, a Christoffel rhythm. Interestingly, Christoffel words always contain a palindrome P surrounded by the first and the last letter of the alphabet. The length of the word, and the number of pulses, is $m + n$, where n counts the 1s, or note onsets, m counts the 0s, or inter-onset pulses; the 0-pulses contain no events. The counts of 0s and 1s are commonly notated as 0^m, 1^n. An important fact is that a word is built from smaller factors, i.e. substrings, so that the concatenation of words is then notated as a multiplication of words (Lothaire 1997).

A Christoffel word can be constructed geometrically on the integer lattice $\mathbb{Z}^2$, see Fig. 6.1. First, one draws a line between the origin and the point (n, m); its slope is $\frac{m}{n}$ (rise over run). Because $m \perp n$, the resulting diagonal does not touch any other points of the lattice. A Christoffel word is a path which follows this diagonal as close as possible. It encodes horizontal steps with the letter a, and vertical ones with b (Berstel 2009; Glen 2012). Figure 6.1 shows the upper path of of $C(3, 5)$, which is $babaabaa$, as well as the lower path $aabaabab$. The strings represent the upper and lower Christoffel word. The only difference between the two is an exchange in the position of the first and the last letter. The reverse of the lower string is the upper string. For the rest of this chapter we will always refer to the the *lower* Christoffel word with the notation $C(m, n)$, see Algorithm 1 for its computation. We will also show how a particular bit operation relates $C(3, 5)$ to its dual $C(5, 3)$ so that the latter becomes the reverse string of the former.

Algorithm 1 Algorithm for Creating Lower Christoffel Words.

```
1: procedure ChristoffelWord(m, n)                    ▷ m, n ∈ N
2:     prev ← 0
3:     i ← 1
4:     output ← empty string vector
5:     while i <= n do
6:         y ← ⌊i m/n⌋
7:         test ← y − prev
8:         output.append 'a'
9:         if test > 0 then
10:            k ← 0
11:            while k < test do
12:                output.append 'b'
13:                k ← k + 1
14:            end while
15:        end if
16:        prev ← y
17:        i ← i + 1
18:    end while
19:    return output
20: end procedure
```

The hierarchy of Christoffel words can be represented by the Stern–Brocot tree (Berstel 2009). The ancestors of Christoffel words $C(m, n)$, with $m < n$, can be obtained by selecting the right and left neighbour of the integer ratio m/n from the Farey sequence F_n, with $n \in N$. A Farey Sequence F_n is the ordered list of all integer ratios in their lowest terms between 0 and 1, and whose denominators do not exceed n (Hardy and Wright 2008). We will discuss Farey Sequences in detail in the next chapter. Every ratio $\in F_n$, except the first $(0/1)$ and the last $(1/1)$, are the result of building the median fraction from its neighbouring ratios that are also in F_{n-1}. This triple of ratios is closest together to $0/1$ and $1/1$ within the Stern–Brocot tree. Therefore, beginning with any Christoffel word, it is possible to trace back all Christoffel words that are substrings. Example: $C(3, 5) = aabaabab$. The neighbouring ratios of $3/5$ in F_5 are $1/2$, which was formed by $0/1 + 1/1$ in F_1, and $2/3$, which came from $1/2 + 1/1$ in F_2. In total, $C(3, 5)$ contains the Christoffel words $C(0, 1)$, $C(1, 1)$, $C(1, 2)$ and $C(2, 3)$. In other words, the strings a, ab, aab and $aabab$ are all those substrings of $aabaabab$ that are also Christoffel words, respectively. The ancestors of Christoffel words $C(m, n)$, where $m > n$, can be obtained in the same way after swapping m and n in order to trace back $C(n, m)$ and its ancestors. As a result, one finds that the duals of the ancestors of $C(n, m)$ are the ancestors of $C(m, n)$.

In combinatorics on words, one studies, for example, the possibilities to rotate a word (Lothaire 1997). For instance, the Christoffel word $C(1, 3) = aaab$ has four rotations; in lexicographic order they are: $aaab \prec aaba \prec abaa \prec baaa$. A word that is strictly smaller in lexicographic order than all of its non-trivial rotations (i.e. conjugates) is called a Lyndon word (Berstel and Perrin 2007). In our example, $C(1, 3) = aaab$ is a Lyndon word, but not its rotational shift $aaba$. Christoffel words are a subset of Lyndon words. In addition, there exist words that are strings of non-negative integers, which produce their own conjugates from *increasing* the first character by one and *decreasing* the last character by one. These words are called Euclidean strings. Binary Euclidean strings are Christoffel words over the alphabet $\{0, 1\}$ (Ellis et al. 2003), a fact that relates Christoffel words also to Sturmian words and Fibonacci strings. Lyndon words and Christoffel words are aperiodic necklaces (Sawada and Williams 2013); for example, $C(3, 4)$ is $aababab$. The dual $C(4, 3)$ results not in a rotation, but in a different word: $abababb$. The reverse of the lower Christoffel word is equal to increasing the first letter and decreasing the last. It results in the upper Christoffel word and this fact is a shared property with Euclidean strings. All binary Euclidean strings are Christoffel words.

6.2.1 Operations on Christoffel Rhythms

In addition to the mapping of $C(m, n)$ to note onsets, three bitwise operations were selected to generate a total of four different rhythmic patterns: 1. $C(m, n)$ over the two-word alphabet $\{a, b\}$ and the morphism $a \mapsto 1, b \mapsto 0$, this is the original Christoffel rhythm. 2. $\neg(C(m, n))$, the binary complement of the previous pattern.

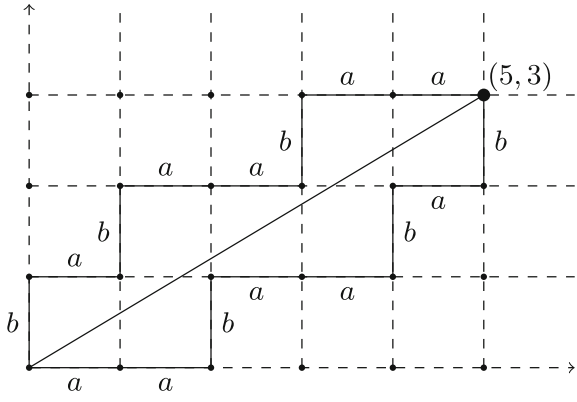


Fig. 6.1 Geometric construction of the upper and lower Christoffel word $C(3, 5)$ with slope $\frac{3}{5}$

3. $C(m, n) \vee (\neg(C(m, n))$ with one additional rule; a repetition of a bit in either $C(m, n)$ or $\neg(C(m, n))$ results in a *zero* bit. This operation effectively marks all changes between consecutive bits as *one*, and repetitions of bits are set to *zero*. The notation to represent this operation is $\nu(C(m, n))$. 4. the complement of the string resulting from previous operation: $\neg(\nu(C(m, n)))$. This operation sets every repeating bit to 'one', the remaining ones are set to 'zero'. As an example, the lower Christoffel word of Fig. 6.1 leads to the following Christoffel rhythms:

$$
\begin{aligned}
C(3, 5) &= \{1, 1, 0, 1, 1, 0, 1, 0\} \\
\neg(C(3, 5)) &= \{0, 0, 1, 0, 0, 1, 0, 1\} \\
\nu(C(3, 5)) &= \{1, 0, 1, 1, 0, 1, 1, 1\} \\
\neg(\nu(C(3, 5))) &= \{0, 1, 0, 0, 1, 0, 0, 0\}
\end{aligned}
$$

The reverse rhythm, notated $R(C(3, 5))$, is $\{0, 1, 0, 1, 1, 0, 1, 1\}$.

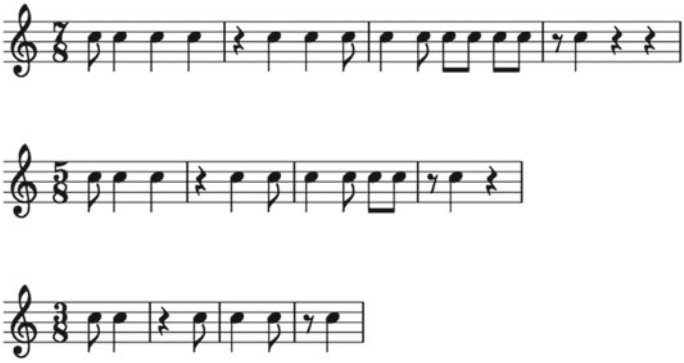



Fig. 6.2 Bit operations on $C(3, 4)$ and its ancestors $C(2, 3)$ and $C(1, 2)$. Each line starts with the original Christoffel word, with each bar representing one of the four operations

The reverse of the complement of the dual is a Christoffel rhythm:

$$R(\neg(C(n,m)) = C(m,n)$$

Figure 6.2 shows the operations working with the rhythm $C(3,4)$ and its ancestors $C(2,3)$ and $C(1,2)$.

6.3 The Burrows-Wheeler Transform as a Tool for Rhythm Analysis

In this section, we investigate the use of our new tool for rhythm analysis: The Burrows-Wheeler Transform (BWT). Remember from Chap. 5, that the BWT of a word w arranges all of its non-trivial conjugates into a matrix in lexicographic order (Berstel 2009). The last column of the matrix is the transform $BWT(w)$. For a Christoffel word $C(m,n)$ over the alphabet $\{a,b\}$, the transform has the form $BWT(C(m,n)) = b^m a^n$. And, because Christoffel words are Lyndon words, the first and the last row of the matrix are the lower and upper Christoffel word, respectively. Furthermore, any two consecutive rows of the matrix differ in exactly two consecutive positions (Berstel 2009), see Fig. 6.4.

The inverse BWT (iBWT) of a lower Christoffel word finds it as the first word, i.e. the first row, of the finished matrix. It is characteristic for the iBWT to visit all factors, i.e. substrings, of the circular word in lexicographic order. This is most interesting from the musical point of view because the iBWT process creates a systematic overview of the combinatorial changes for any given rhythmic pattern, which helps to gain a deeper knowledge of the compositional material. Progressive shortening or expansion of patterns are key concepts in classical music theory (Caplin 2000), and we have seen examples of this technique when applied to rhythmic patterns in Chap. 4. The BWT is a tool for analysis that allows us to see relationships between different patterns, including also patterns that have been rotated. If a rhythm pattern p is a substring of some Christoffel word, we denote it as $p \subset C(m,n)$. To identify the substring's position within the matrices of the iBWT,

$$\tau_j^i(C(m,n))$$

denotes its lexicographic position i within the matrix of rows with length j. Note that $\tau_{m+n}^0(C(m,n)) = C(m,n)$.

A conjugate, or rotation of a word, is notated as $\rho^i(C(m,n))$, with i denoting the index into the lexicographically ordered list of all rotations, and with $\rho^0(C(m,n)) = C(m,n)$. For example: $\rho^7(C(5,3)) = \{0,1,0,1,1,0,1,1\}$. Note that the lexicographically largest rotation of a lower Christoffel word is identical with the upper Christoffel word.

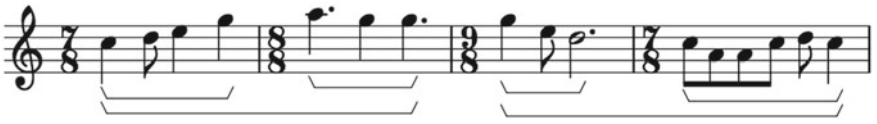



Fig. 6.3 A musical sentence with two phrases containing two patterns each. The patterns have the Christoffel rhythms: $\rho^1(C(3,4))$, $\rho^1(C(5,3))$, $\rho^1(C(1,2)) \cdot C(5,1)$, $C(1,6)$

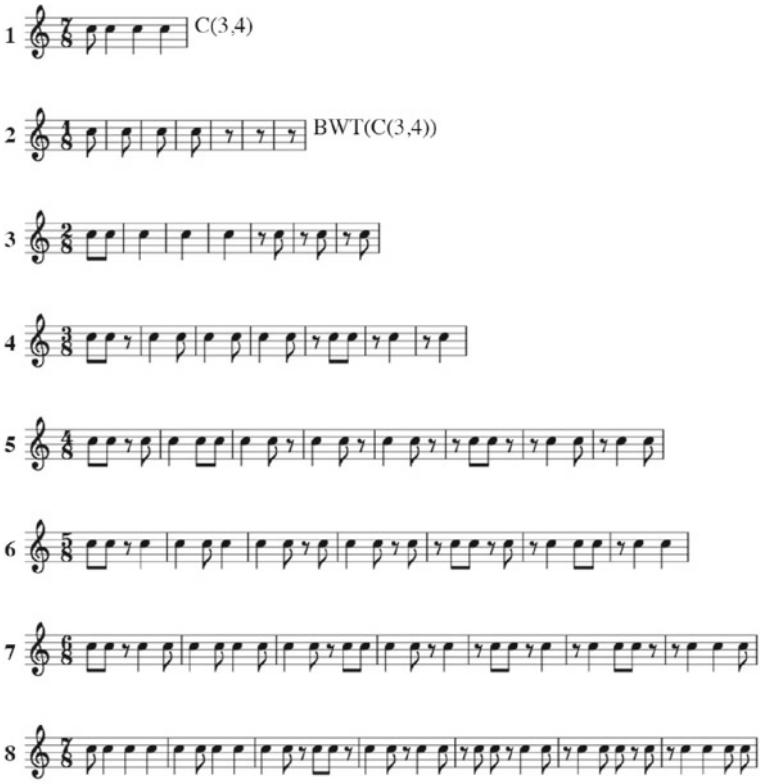



Fig. 6.4 Inverse BWT matrix of $C(3,4)$ in rows 2–8. The second row is the BWT itself, the last row shows all rotations of $C(3,4)$ in lexicographic order. Rows 3–7 show rotations of substrings, *if and only if* the rotation is a substring of $C(3,4)$

We will now demonstrate the usefulness of the BWT for music composition and analysis. Figure 6.3 shows a 31-pulse long rhythm sentence composed out of two phrases, which split up into two or three patterns each. The rhythm transcribed into SNMR reads `>II -I- >H~I::X` For the demonstration of the BWT we concentrate on the very first 7-pulse pattern. This pattern is $\rho^1(C(3,4)) = \{1,0,1,1,0,1,0\}$. Figure 6.4 shows the matrices of the iBWT of $C(3,4)$, which

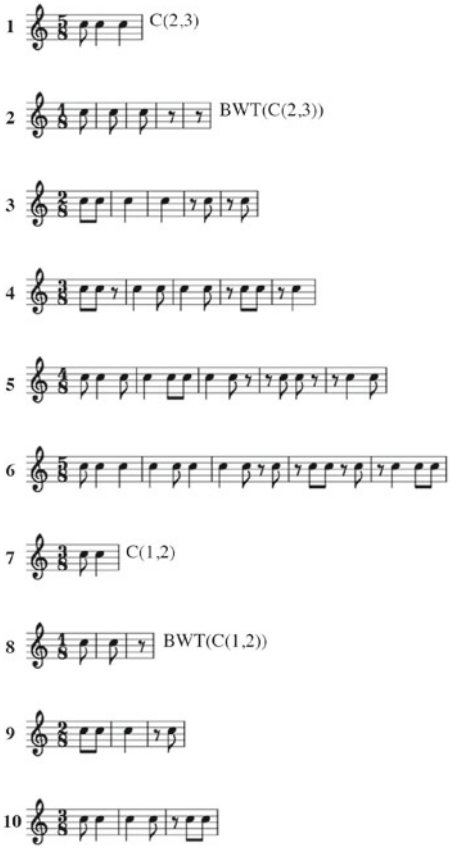



Fig. 6.5 Inverse BWT matrix of the ancestors of $C(3, 4)$, which are $C(2, 3)$ (rows 1–6), and $C(1, 2)$ (rows 7–10). The rows 6 and 10 show all rotations of the words in lexicographic order

returns all possible substrings with lengths 1 to $m + n$ in lexicographic order. Figure 6.5 shows the iBWT of the ancestors of $C(3, 4)$, which are $C(2, 3)$ and $C(1, 2)$. We ignore $C(0, 1)$ and $C(1, 1)$ because they are not useful in a musical context.

6.4 Rhythms from Various Music Cultures

In this section, we investigate the relationship of Christoffel words with rhythmic patterns that are used in various musical cultures. The idea here is to demonstrate the characteristics of these rhythms, and to show that Christoffel rhythms can be used to represent them.

Table 6.1 Greek verse rhythms and their representation as Christoffel rhythms

Foot	Christoffel rhythm
Iamb X	C(1,2) $R(\neg(C(2,1)))$ $\nu(C(2,1))$
Trochee >	$\rho(C(1,2))$ $\nu(C(1,2))$ $R(\nu(C(1,2)))$
Tribach i	$\tau_3^0(C(1,3)$ $\tau_3^0(C(1,4))$
Spondee II	C(2,2) $R(\neg(C(2,2)))$
Anapest :I	$C(1,3)$ $\tau_4^0(C(2,5))$ $R(\neg(C(3,1)))$
Dactyl I:	$\rho^2(C(1,3))$ $\tau_4^1(C(2,3))$ $\tau_4^1(C(3,4))$
Procleusmatic ::	$\tau_4^0(C(1,4))$ $\tau_4^0(C(1,5))$ $\nu(C(2,2))$ $R(\nu(C(2,2)))$
Amphibrach .I.	ρ(C(1,3)) $\subset$ C(2,3) $\subset$ C(2,4)

Greek Verse Rhythms

The verse rhythms of the ancient Greeks have been used often in the analysis of Western Classical music (Dupré 1925; Georgiades 1982; Markevitch 1983; Messiaen 1995). See also Chap. 3 on shorthand notation, Sect. 3.4.3. One of the reasons for their popularity is the fact that these rhythms encode small cells with a progressive number of onsets that are applicable to many different forms and styles in music. Not only is the relationship of our Western culture with Greek poetry and philosophy an important one, also the fact that these verse rhythms occur frequently as archetypes throughout the history of music is astonishing. It was only natural to investigate, if there was any relationship with mathematical principles. The Tables 6.1 and 6.2 display all Greek verse Rhythms and their representations as Christoffel rhythms. Here, one finds that Greek verse patterns are represented by Christoffel words and their rotations, or by their subsets. Only two verse rhythms are products of Christoffel words. Most rhythms can have more than one explanation, in which case the shortest solution is preferred. The range of n is $\{1, 2, 3, 4, 5\}$, the possible integers for m are $\{1, 2, 3, 4, 6\}$. It is common in Classical Greek poetry to use variations of a single verse rhythm (Silk 2004). It happens also frequently that combinations of different verse rhythms are being used within the same piece of poetry.

Table 6.2 Greek verse rhythms with five or more pulses

Rhythmic pattern	Christoffel rhythm
XI Bacchius	$C(2,3)$; $R(\neg(C(3,2)))$
>I Amphimacer	$\rho^1(C(2,3))$
I> Antibacchius	$\rho^2(C(2,3))$
>: Peon 1	$\rho^3(C(1,4))$; $R(\nu(C(2,3)))$
X: Peon 2	$\rho^2 C(1,4)$
:> Peon 3	$\rho^1(C(1,4))$; $R(\nu(C(2,3)))$
:X Peon 4	$C(1,4)$; $\nu(C(3,2))$; $R(\neg C(4,1))$
XX Diiamb	$C(2,4)$
>> Ditrochee	$\nu(C(2,4))$; $R(\nu(C(2,4)))$; $\rho^2(C(2,4))$
I:I Choriamb	$\tau_6^3(C(2,5))$
X> Antispast	$\tau_6^0 C(3,4)$
II: Ionic Major	$\tau_6^2 C(3,4)$
:II Ionic Minor	$C(1,3) \cdot C(1,1)$
III Molossus	$C(3,3)$; $R(\neg(C(3,3)))$
XII Epitrite 1	$C(3,4)$; $R(\neg(C(4,3)))$
>II Epitrite 2	$\rho^1(C(3,4))$
I>I Epitrite 3	$\rho^2(C(3,4))$
II> Epitrite 4	$\rho^3(C(3,4))$
X>I Dochmius	$\rho^1(C(3,5))$
IIII Dispondee	$C(4,4)$; $R(\neg C(4,4))$
>:III Dactylo-Epitrite	$\rho(C(1,3)) \cdot C(1,2) \cdot C(2,2)$
-IIII	$\rho^4(C(6,5))$

Latin-American Rhythms

The bell-patterns and claves of Latin-American rhythm can be regarded as transformations of the African musical traditions; coming from the same African people who were enslaved and displaced into the West-Indies and America (Giger 1993). Therefore, I would like to start the survey with two famous bell-patterns from Africa. The Ewe rhythm from Ghana is a 12-pulse pattern that can be found in its rotated forms also in other African music cultures. The Ewe rhythm has the duration pattern $\{2,2,1,2,2,2,1\}_{12}$, IIXI> which is a conjugate of $C(5,7)$, namely $\rho^5(C(5,7))$. $\{5,7\}$ is the only coprime pair of positive integers that forms a partition of 12. As a seven beat sequence, the Ewe rhythm is maximally even (London 2012). Another prominent rhythm is the 16-pulse Ghana bell-pattern $\{3,3,4,4,2\}_{16}$, --HHI. It can be represented by the product of $C(7,3) \cdot \tau_6^1(C(5,2))$.

African bell-patterns have clearly influenced the musical styles of Latin America. Table 6.3 shows the most important bell-patterns and clave rhythms used in Afro-Cuban styles and traditions, which also had their influence on Brazilian Samba music.

Table 6.3 Latin-American and African bell-patterns and claves

Name	Pattern	Christoffel rhythm
3-2 claves	`--HI H`	$\tau_{12}^{1}(C(9,4)) \cdot C(3,1)$
	`-H---`	$\rho^{4}(C(11,5))$
	`-- .- I H`	$C(4,2) \cdot \nu(C(3,1)) \cdot C(1,1) \cdot C(3,1)$
	`-- .---`	$C(4,2) \cdot R(\neg(\nu(C(7,3))))$
3-2 Rhumba	`-H- I H`	$\rho^{1}(C(7,3)) \cdot C(1,1) \cdot C(3,1)$
Bossa Nova	`--H --`	$\rho^{2}(C(11,5))$
Soukous	`--H X^I`	$C(7,3) \cdot \nu(C(5,1))$
Gahu	`--H HI`	$C(7,3) \cdot \tau_{6}^{1}(C(5,2))$
Tresillo	`--I`	$\rho^{2}(C(5,3))$
Cinquillo	`IXX`	$\rho^{2}(C(3,5))$
Baqueto of Danzon	`IXX IIII`	$\rho^{2}(C(3,5)) \cdot C(4,4)$
6/8 clave	`II X>I`	$\rho^{4}(C(5,7))$
Brazilian Samba	`II >II >I`	$\rho^{6}(C(7,9))$

The Montuno is a Cuban 16-pulse rhythm pattern played by the piano as part of the rhythm section (Mauleon 2011). It has many variations, but the basic structure is $\{2,1,2,2,2,2,2,1\}_{16}$, `>II II I>`. The corresponding Christoffel rhythm is $\tau_{16}^{1}(C(8,9))$. The cycle of rotations of $\tau_{16}^{k}C(8,9)$ is musically most interesting: Two ternary chunks are in contrast with five binary chunks. This is one of the two partitions of 16 that contain only 2s and 3s as parts.[2]

The Bolero rhythm $\{2,1,1,2,2,2,2,2,2\}_{16}$, `I:II IIII` has the composition $\tau_{6}^{3}(C(2,5)) \cdot C(5,5)$. The Merengue $\{3,1,2,2,2,2,1,1,1,1\}_{16}$, `-.II II:` has a pattern of accents resulting in the 3-2 son clave, `--HHI`. The rhythm is composed out of three Christoffel rhythms: $\nu(C(1,3)) \cdot C(4,4) \cdot \nu(C(2,2))$. Its tone colours change in the following way: *aababaaabb*, where $a \mapsto high$ and $b \mapsto low$. The Christoffel rhythm of that colour change is $\tau_{6}^{0}(C(3,4)) \cdot \nu(C(3,1))$.

This is one of many possible cencerro (cowbell) rhythms (Giger 1993, p. 214):

$$\{2,2,1,1,1,2,1,1,2,1,1,1\}_{16}.$$

`IIiI:Ii`

The corresponding Christoffel rhythm is $C(1,1) \cdot \rho^{9}(C(3,11))$. The change of tone colours is *abbbabbabba*, which corresponds to $\tau_{11}^{3}(C(9,4))$. Tone colours and durations of the rhythm change independently. This is a typical African way of creating polyrhythm (Arom 1991).

[2]The other partition has four ternary and and two binary chunks, and it can be found in Jerry Gray's Swing classic *A String Of Pearls*: 3, 3, 3, 3, 2, 2.

Table 6.4 Turkish patterns

Name	Pattern	Christoffel rhythm
Düyek	`.I. II`	$C(3,5)$
Süreyya	`I>`	$\rho^2(C(2,3))$
Devri hindi	`iII`	$C(1,4)\cdot C(1,1)$
Mandira	`>.I.`	$\rho^3(C(2,5))$
Katakofti	`-I-`	$\rho^1(C(5,3))$
Çifte düyek	`IHI III:`	$C(1,1)\cdot C(3,1)\cdot C(3,3)\cdot \rho^2(C(1,3)$
Aksak	`I:I>`	$\rho^2(C(1,3))\cdot \rho^2(C(2,3))$
Aksak semai	`>II>`	$\rho^1(C(2,3))\cdot \rho^2(C(2,3))$
Sofiyan	`I:`	$\rho^2(C(1,3))$
Sofiyan 2	`HI>`	$C(3,1)\cdot \rho^2(C(2,3))$
Evfer (mevievi)	`I:>I`	$\tau_9^4(C(3,7))$
Devri revan.	`-II -II`	$\rho^4(C(8,6))$
Fahte	`I:II IIH ::`	$\tau_8^8(C(15,22))\cdot C(2,2)\cdot C(3,1)\cdot \tau_8^8(C(1,4))$
Lenk fahte	`I-X:`	$C(3,2)\cdot \rho^2 C(1,4)$
Devri kebir	`HHH II:I`	$C(9,3)\cdot \tau_8^{15}(C(15,22))\cdot$
	`WW H`	$\tau_{20}^2(C(20,3))\cdot$
	`W IIII`	$C(7,1)\cdot C(4,4)$

Turkish Rhythms

The rhythms used in Turkish classical music are shown in Table 6.4 (Giger 1993). Amongst these traditional patterns features a particular rhythm called *aksak*, which literally means 'limping'. Aksak is also the name for a whole family of rhythms, which have been analyzed in great detail by ethnomusicologists (Arom 2004; Brailoiu 1951). We mentioned the Aksak theory of both authors earlier in Chap. 4, Sect. 4.7.1, within the context of the seven categories of rhythmic patterns.

Macedonian Rhythms

The dance rhythms of the folk music of Macedonia combine together beats of different lengths, usually two or three pulses long (Giger 1993), see Table 6.5. They all correspond to single Christoffel rhythms and their rotations, except for the last 14-pulse long rhythm, which is a product. The individual beats of the patterns can be further subdivided into the smallest unit, e.g. `-II`, might become `>II`, or `>:I`. In this way, one can develop many rhythmic and melodic variations.

Rhythmic Oddity

Rhythmic patterns that follow the principle of *rhythmic oddity* were discovered by Arom and a computational method using Lyndon words is described by Chemiller and Truchet (2003). The definition of rhythmic oddity means that a pattern based on an even number of pulses cannot be broken into two halves of equal length, no matter which breaking point is chosen within the circular pattern. Those kinds of rhythms

Table 6.5 Macedonian patterns

Pattern	Christoffel rhythm
`I-`	$C(3,2);\ R(\neg(C(2,3)))$
`-II`	$\rho^2(C(4,3))$
`II-`	$C(4,3);\ R(\neg(C(3,4)))$
`-I-`	$\rho^7(C(5,3))$
`III-`	$C(5,4);\ R(\neg(C(4,5)))$
`-II-`	$\tau_{10}^3 C(7,5)$
`II-II`	$\rho^2(C(6,5))$
`I-III`	$\rho^3(C(6,5))$
`II-I-`	$C(7,5);\ R(\neg(C(5,7)))$
`I-II-`	$\rho^1(C(7,5))$
`-I-I-`	$\rho^2(C(8,5))$
`-I--I`	$\rho^3(C(8,5))$
`I-III-`	$\tau_{14}^3 C(9,7)$
`III-I-`	$C(5,4)\cdot C(3,2)$

Table 6.6 Some African patterns, which show the principle of rhythmic oddity

Rhythmic pattern	Christoffel rhythm
`--I`	$\rho^2(C(5,3))$
`-I-II`	$\rho^4(C(7,5))$
`-II-III`	$\rho^6(C(9,7))$
`-III-IIII`	$\rho^8(C(11,9))$
`-IIII -IIIII`	$\rho^{10}(C(13,11))$
`----I`	$\rho^4(C(9,5))$
`--I--II`	$\tau_{18}^4(C(12,7))$
`--I-I-I`	$\rho^6(C(11,7))$
`--II--III`	$\rho^2(C(5,3))\cdot C(5,3)\cdot C(3,3)$
`--II-I-II`	$\rho^2(C(5,3))\cdot C(6,4)\cdot C(2,2)$
`-I-I-I-II`	$\rho^8(C(13,9))$
`--- ---I`	$\rho^6(C(13,7))$
`---I ---II`	$\tau_{24}^6 C(16,9)$
`---I --I-I`	$\tau_{24}^9 C(17,10)$
`--I --I --I`	$\rho^6(C(15,9))$

feature prominently amongst certain African music traditions. People like the Aka, Gbaya, Ngbaka, Nzakara, or the Zande are using those rhythms. Table 6.6 shows the patterns listed by Chemiller and Truchet, together with the corresponding Christoffel rhythm. It is a characteristic of these patterns that they are based on combinations of 2s and 3s. Nine rhythms are conjugates of Christoffel words, three are subsets of $C(m,n)$, and two of them a constructed by a product of Christoffel words and their rotations.

6.4.1 Euclidean Rhythms

Euclidean rhythms are patterns of note onsets and silences that are distributed as evenly as possible over n pulses (Demaine et al. 2009). Euclidean rhythms are generated by the Bjorklund algorithm that was developed for the control of repetition rates and evenly spaced timing patterns in spallation neutron source (SNS) accelerators (Bjorklund 2003). Toussaint found a most interesting connection between those evenly distributed timing patterns and rhythmic ostinato patterns used around the globe by many different cultures. All of the published Euclidean Rhythms (Demaine et al. 2009) have a correspondent Christoffel word over the alphabet a, b with $a \prec b$ and the subsequent mapping $a \mapsto 1, b \mapsto 0$. For example the Euclidean rhythm $E(3, 7)$, i.e. three onsets within a seven pulse sequence, $\{1, 0, 1, 0, 1, 0, 0\}$, is identical to the Christoffel word $C(4, 3)$. Note the correspondence between $C(m, n)$ and $E(n, m + n)$. However, some Euclidean rhythms are *rotations* (conjugates) of Christoffel words, for example $E(4, 11)$ is $\{1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0\}$, which is a rotation of $C(7, 4)$, or $\{1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0\}$. Our tests show that all published Euclidean rhythms (Demaine et al. 2009), and all patterns generated by the Bjorklund algorithm, are either Christoffel words or one of their rotations. We can present five exclusive conditions for equality: First, if the slope $\frac{m}{n}$ of $C(m, n)$ is a super-particular ratio, i.e. if $m - kn = 1$, with $k \in N$, then $C(m, n) = E(n, m + n)$. Second, it is also true for Christoffel words with the slope $\frac{1}{n}$, and third, for Christoffel words with the slope m. If, however, the slope $\frac{m}{n}$ is not a super-particular ratio, then there are two remaining possibilities for a mapping: Either the nth rotation of $C(m, n)$ in lexicographic order is identical with $E(n, m + n)$, or, and this is the fifth condition, by shifting the proper suffix of $C(m, n)$ to the beginning of the word, one obtains $E(n, m + n)$. By proper suffix we mean the very last factor of the Christoffel word, which is also a Christoffel word of the form $a^x b^y$. Here is an example for the relationship between the nth conjugate of $C(m, n)$ in lexicographic order, and $E(n, m + n)$: Consider the complete list of conjugates of $C(7, 5)$ in lexicographic order in Table 6.7, which can be produced as a by-product of the iBWT.

The fourth conjugate, $n - 1$, is the Euclidean rhythm $E(5, 12)$ with the mapping $a \mapsto 1, b \mapsto 0$. Finally, the next example shows how a particular Christoffel word can be transformed into an Euclidean rhythm by shifting its proper suffix to the front: Consider the Christoffel word $C(8, 5) = ababbababb\mathbf{abb}$. Its last factor, the proper suffix **abb**, is the Christoffel word $C(2, 1)$. By shifting it to the front we obtain $E(5, 13) = \mathbf{abb}ababbababb$. Tables 6.8 and 6.9 show all published Euclidean rhythms (Demaine et al. 2009) and their corresponding Christoffel word together with a classification of their relationship according to the principles explained above. Based on these tables and on further tests with the Bjorklund algorithm we can confirm that these are the only relationships known so far.

Table 6.7 Conjugates of $C(7, 5)$ in lexicographic order. The $(n-1)^{\text{th}}$ item is a Euclidean rhythm

`abababbabab`	0	$C(7, 5)$
`ababbababab`	1	
`ababbababbab`	2	
`abbababababb`	3	
abbababbabab	4	$E(5, 12)$
`bababababbab`	5	
`bababbababab`	6	
`babababbabba`	7	
`babbababbaba`	8	
`babbababbaba`	9	
`bbababababba`	10	
`bbababbababa`	11	

Table 6.8 The first three known conditions for equality between a Euclidean rhythm $E(n, m+n)$ and a Christoffel word $C(m, n)$. The other two conditions are shown in Table 6.9

Euclidean rhythm	Christoffel word
$\frac{m}{n}$ is a super-particular ratio	
$E(2, 5)$	$C(3, 2)$
$E(3, 7)$	$C(4, 3)$
$E(4, 9)$	$C(5, 4)$
$E(5, 11)$	$C(6, 5)$
$E(6, 13)$	$C(7, 6)$
$E(7, 15)$	$C(8, 7)$
$E(8, 17)$	$C(9, 8)$
$E(5, 16)$	$C(11, 5)$
$\frac{m}{n} = \frac{1}{n}$	
$E(2, 3)$	$C(1, 2)$
$E(3, 4)$	$C(1, 3)$
$E(4, 5)$	$C(1, 4)$
$E(5, 6)$	$C(1, 5)$
$E(6, 7)$	$C(1, 6)$
$E(7, 8)$	$C(1, 7)$
$\frac{m}{n} = m$	
$E(1, 3)$	$C(2, 1)$
$E(1, 4)$	$C(3, 1)$
$E(1, 5)$	$C(4, 1)$

Table 6.9 Two additional conditions for equality between a Euclidean rhythm $E(n, m+n)$ and a Christoffel word $C(m, n)$. See also Table 6.8

Euclidean rhythm	Christoffel word
$E(n, m+n) = \rho^{n-1}(C(m, n))$	
$E(3, 8)$	$C(5, 3)$
$E(3, 11)$	$C(8, 3)$
$E(3, 14)$	$C(11, 3)$
$E(4, 11)$	$C(7, 4)$
$E(4, 15)$	$C(11, 4)$
$E(5, 7)$	$C(2, 5)$
$E(5, 9)$	$C(4, 5)$
$E(5, 12)$	$C(7, 5)$
$E(7, 9)$	$C(2, 7)$
$E(7, 10)$	$C(3, 7)$
$E(7, 16)$	$C(9, 7)$
$E(7, 17)$	$C(10, 7)$
$E(9, 22)$	$C(13, 9)$
$E(11, 24)$	$C(13, 11)$
Rotation of suffix of $C(m, n)$	
$E(5, 8)$	$C(3, 5)$
$E(5, 13)$	$C(8, 5)$
$E(7, 12)$	$C(5, 7)$
$E(7, 18)$	$C(11, 7)$
$E(8, 19)$	$C(11, 8)$
$E(9, 14)$	$C(5, 9)$
$E(9, 16)$	$C(7, 9)$
$E(9, 23)$	$C(14, 9)$
$E(13, 24)$	$C(11, 13)$
$E(15, 34)$	$C(19, 15)$

6.5 Conclusion

We have presented in this chapter how Christoffel words can generate musical rhythms and meters from many different cultures. Three bitwise operations are able produce unique counterparts to the original form. The application of the inverse Burrows-Wheeler Transform demonstrates that any rhythmic pattern can be reduced to a product of Christoffel words and their conjugates. We have, therefore, introduced a definition for a Christoffel rhythm that includes bitwise and reverse operations as well as the analysis of its factors, i.e. substrings that are Christoffel words themselves.

Christoffel words are closely connected to Euclidean rhythms. Demaine et al. (2009) showed in their paper that only a subset of Euclidean strings are Euclidean rhythms. Ellis et al. (2003) showed that all binary Euclidean strings are Christoffel

words. In addition, we have demonstrated that there are five ways to transform a Christoffel word into a Euclidean rhythm and vice versa. Therefore, it is valid to say that all binary Euclidean strings can be transformed via Christoffel words into Euclidean rhythms.

Christoffel rhythms can be viewed as a sort of DNA for a great variety of rhythms from different cultures. In Western theory one might call them additive or compound rhythms and meters. Each individual word $C(m, n)$ is then like a gene that can be read in various ways: forwards and backwards, in cyclic rotation, and in parts. Rotations and substrings can be computed with the Burrows-Wheeler Transform and its inverse transform. When used as factors, one can concatenate Christoffel rhythms in order to build longer and more complex rhythmic patterns.

We have also presented examples where Christoffel words describe tone colour changes. Rhythmic changes of tone colours constitute another most important rhythmic element, which, when perceived together with the changes in durations, produces a polyphony of musical events. One can, therefore, extend the analysis of changes via Christoffel rhythms to other musical parameters, for example, dynamic changes and accentuation patterns.

For the composer, this approach is equally useful. A single Christoffel word $C(m, n)$ and its iBWT encapsulates a large number of possibilities for rhythmic patterns. For each $C(m, n)$, after calculating its iBWT matrix, the number of patterns equals $(m + n)^2$. Even if one subtracts repetitions of patterns within the iBWT matrix, this still leaves a very large pool of possibilities for musical variation based on a single Christoffel word. The scope for Christoffel rhythms comprises all musical parameters that can be represented rhythmically, such as changes of timbre, articulation, dynamics, register, and note durations.

References

Arom, S. (1991). *African polyphony and polyrhythm*. Cambridge: Cambridge University Press. (ISBN 052124160X).

Arom, S. (2004). L'aksak: Principes et typologie. *Cahiers de musiques traditionnelles*, *17*, 11–48.

Berstel, J. (2009). *Combinatorics on words: Christoffel words and repetitions in words* (Vol. 27). CRM monograph series. Providence: American Mathematical Society. (ISBN 9780821873250).

Berstel, J., & Perrin, D. (2007). The origins of combinatorics on words. *European Journal of Combinatorics*, *28*(3), 996–1022. http://www.sciencedirect.com/science/article/pii/S0195669805001629, https://doi.org/10.1016/j.ejc.2005.07.019. Retrieved March 2018.

Bjorklund, E. (2003). The theory of rep-rate pattern generation in the SNS timing system. Technical report. Los Alamos, N.M.: Los Alamos National Laboratory. (SNS-NOTE-CNTRL-99).

Brailoiu, C. (1951). Le rythme Aksak. *Revue de Musicologie*, *33*(99/100), 71–108. http://www.jstor.org/stable/926002.

Caplin, W. E. (2000). *classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven*. Oxford: Oxford University Press. (ISBN 019514399X).

Chemillier, M., & Truchet, C. (2003). Computation of words satisfying the "rhythmic oddity property" (after Simha Arom's works). *Information Processing Letters*, *86*(5), 255–261. https://doi.org/10.1016/S0020-0190(02)00521-5.

Demaine, E. D., GomezMartin, F., Meijer, H., Rappaport, D., Taslakian, P., Toussaint, G. T., et al. (2009). The distance geometry of music. *Computational Geometry*, *42*(5), 429–454. https://doi.org/10.1016/j.comgeo.2008.04.005.

Dupré, M. (1925). *Traité d'Improvisation á l'Orgue* (Vol. 2). Paris: Alphonse Leduc. (Out of print). http://quod.lib.umich.edu/. Retrieved December 2010.

Ellis, J., Ruskey, F., Sawada, J., & Simpson, J. (2003). Euclidean strings. *Theoretical Computer Science*, *301*(1), 321–340. https://doi.org/10.1016/S0304-3975(02)00589-3.

Georgiades, T. G. (1982). *Music and language: The rise of Western music as exemplified in settings of the mass*. New York; Cambridge: Cambridge University Press.

Giger, P. (1993). *Die Kunst des Rhythmus: Professionelles Know How in Theorie und Praxis*. Mainz, London, New York: Schott. (ISBN 978-3-7957-1862-6).

Glen, A. (2012). Combinatorics of Lyndon words. https://amyglen.files.wordpress.com/. Retrieved August 2015.

Hardy, G., & Wright, E. (2008). *An introduction to the theory of numbers* (6th ed.). Oxford: Oxford University Press.

London, J. (2012). *Hearing in time. Psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Lothaire, M. (1997). *Combinatorics on words*. Cambridge mathematical library. Cambridge: Cambridge University Press. (ISBN 0521599245).

Lothaire, M. (2005). *Applied combinatorics on words* (Vol. 105). Cambridge: Cambridge University Press. https://books.google.ca/books?id=fpLUNkj1T1EC. (ISBN 978-0521848022).

Markevitch, I. (1983). *Die Sinfonien von Ludwig van Beethoven. Historische, analytische und praktische Studien*. Leipzig, GDR: Edition Peters.

Mauleon, R. (2011). *The Salsa Guidebook*. Sher Music. (ISBN 978-0961470197).

Messiaen, O. (1995). *Traité de Rhythme, de Couleur et d'Ornithologie* (Vol. 1). Alphonse Leduc. (ISBN 978-004628892).

Sawada, J., & Williams, A. (2013). A gray code for fixed-density necklaces and Lyndon words in constant amortized time. *Theoretical Computer Science*, *502*, 46–54.

Silk, M. S. (2004). *Homer, the Iliad*. Cambridge, New York: Cambridge University Press.

Toussaint, G. (2004). A comparison of rhythmic similarity measures. Technical report, School of Computer Science, McGill University. http://cgm.cs.mcgill.ca/~godfried/publications/similarity.pdf (SOCS-TR-2004.6). Retrieved March 2018.

Chapter 7
The Farey Sequence as a Model for Musical Rhythm and Meter

7.1 Introduction

What we have seen so far are rhythmic structures that are based on an underlying small pulsation. We have performed the analysis of asymmetric musical sentences (Sect. 4.4). I have also given a method to compose rhythmic sentences that uses integer partitions (Sect. 4.7). In this context, seven categories of metric grouping patterns have been found useful (Sect. 4.7.1). I have introduced a measure for the musical impact that occurs when a metric group of two pulses changes to a group of three pulses, and vice versa (Sect. 4.5). In the previous Chap. 6, we have learned about the shared properties of patterns from different music cultures when they are being analyzed with Christoffel rhythms and the Burrows-Wheeler transform, see also Chap. 5. Two models of rhythm have been introduced. One that uses integer partitions and metric grouping, the other one takes advantage of the relatively new field of combinatorics on words.

In this chapter, we will see that musical rhythms and metrical hierarchies can be represented by integer ratios. In fact, Western notation practice is to a large extent based on this principle. Note events fit into this framework by alignment with discrete onset times on a metrical grid structure. This metrical grid and the rhythms can be described by filtered Farey Sequences.

We mention a theory, which regards rhythm and meter as a process of entrainment. This theory looks at modes and resonances caused by a rhythmic pattern, which acts like a driving force. The entrained rhythmic patterns, which we have learned from early childhood onwards, will resonate to any close match with a musical rhythm that we might perceive right now. One has discovered mode locking regions in models of meter perception that correlate with a tree of small integer ratios called the *Farey tree* (Large and Kolen 1995). Also, one has found resonant frequencies in models of neural oscillators that show peak amplitudes at small integer ratios between the natural frequency of the oscillator and frequencies present in the rhythmic stimulus (Large et al. 2010). Therefore, the aim of this chapter is to develop a model of musical

rhythm and meter, which looks at the relations between the timing of note events. One of the roots of this theory, which we will discuss first, is the Farey sequence.

Finally, we will take into account important aspects of the human perception of rhythm and meter.

7.2 The Farey Sequence

Farey sequences are a part of number theory, but they are also known to be used in the area of physics (Tomás 2014), where they have been related to the tuning of resonance diagrams in particle accelerators. They can be useful in electronic resistor networks (Khan 2012), and in digital image processing (Pratihar and Bhowmick 2017). In music, Farey sequences appear in James Tenney's *Spectral CANON for CONLON Nancarrow*, see Wannamaker (2012) for a detailed analysis.

A Farey sequence F_n can represent the onset times and relative durations of all metric subdivisions of a beat. In this case, the onset times are marked by fractions between 0 and 1, which is the normalized duration of the beat.

As stated earlier, a fraction a/b belongs to F_n, if

$$0 <= a <= b <= n \tag{7.1}$$

with $a, b, n \in \mathbb{N}$. n is called the order of the Farey sequence. Starting with $F_1 = \{\frac{0}{1}, \frac{1}{1}\}$, one can construct the next higher Farey sequence by inserting mediant fractions between pairs of consecutive terms and by satisfying the above Eq. 7.1. A mediant fraction c/d between a/b and e/f has the form:

$$c/d = (a + e)/(b + f) \tag{7.2}$$

For example, to build F_2 one has to insert $\frac{1}{2}$ into F_1. The first few examples out of an infinite number of F_n are then as follows:

$$\begin{aligned}
F_1 &= \{\tfrac{0}{1}, \tfrac{1}{1}\} \\
F_2 &= \{\tfrac{0}{1}, \tfrac{1}{2}, \tfrac{1}{1}\} \\
F_3 &= \{\tfrac{0}{1}, \tfrac{1}{3}, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{1}{1}\} \\
F_4 &= \{\tfrac{0}{1}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{1}{1}\} \\
F_5 &= \{\tfrac{0}{1}, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}, \tfrac{1}{1}\} \\
F_6 &= \{\tfrac{0}{1}, \tfrac{1}{6}, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}, \tfrac{5}{6}, \tfrac{1}{1}\}
\end{aligned}$$

A Farey sequence F_n has the following important properties:

1. If a/b and c/d are two successive terms of F_n, then $bc - ad = 1$.
2. If a/b, c/d, e/f are three successive terms of F_n, then $c/d = (a + e)/(b + f)$. c/d is called the mediant fraction between a/b and e/f.
3. If $n > 1$, then no two successive terms of F_n have the same denominator.

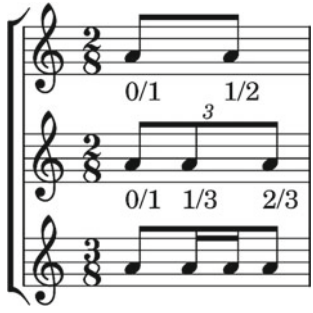



Fig. 7.1 Rhythm of two eighth notes against three. Onsets are marked by elements of F_3. The combined rhythm on the third line is the union of the two upper voices

Next, we see that F_n can represent all possible beat subdivisions. If a quarter note has the length 1, then the onsets of eighth notes are $\{\frac{0}{1}, \frac{1}{2}\}$. The onsets of an eighth note triplet are $\{\frac{0}{1}, \frac{1}{3}, \frac{2}{3}\}$. Uniting both sets and adding $\{\frac{1}{1}\}$ gives us F_3. The fraction $\frac{1}{1}$ marks the onset of the following quarter note. Playing F_3 as a musical rhythm with two voices enables us to listen to the simultaneous rhythm of two eighths plus a three eighths triplet. We see how the Farey sequence F_3 contains all onsets of this polyrhythm in order. Figure 7.1 shows this example written as two voices. The durations of both voices have the ratio $2 : 3$. The third staff shows the notation of this rhythm as a compound polyrhythm. The next highest subdivision would be in four. The fractions to mark their onsets are part of F_4. But F_4 already comprises F_3 and F_2, which contain $\frac{1}{2}$. Therefore, and because all fractions must be in their lowest terms, only the set $\{\frac{1}{4}, \frac{3}{4}\}$ is added to F_3. This results in F_4 containing the onsets of the subdivisions $2 : 3 : 4$ of a quarter note with duration 1. One can continue to add higher subdivisions; F_n contains the subdivision n of a reference duration and all subdivisions m, where $1 < m <= n$, with $m, n \in \mathbb{N}$. Therefore, F_n is a unified representation of the onsets of a polyrhythm. The integer ratios of the durations built from two sets of subdivisions are $a : b$, with $0 < a < b <= n$.

The musical rendering of F_n as a polyrhythm reveals interesting structural properties. The inter-onset intervals (IOIs) between the onsets of the polyrhythms are the differences between consecutive pairs of fractions of F_n. Farey sequences are symmetrical and self-similar sets of fractions. Rhythmically, the second half of a Farey sequence is the mirror of the first half. In musical terms, the onsets in the range of [0.5...1.0] form the *retrograde* of the first half of the Farey Sequence between 0 and 0.5. Retrograde pitch and rhythmic structures have been used widely in Western music. They emerge from a set of transformations that is very common for polyphonic styles of music ranging from Renaissance to the Second Viennese School and beyond. A theme or a series of notes can be subjected to any of these transformations: Transposition, retrograde motion of pitches and rhythms, the inversion of musical intervals, and the permutation of notes and durations, plus a combination of these techniques. The Farey sequence can encompass more than the subdivision of a single beat. It can span over an entire rhythmic phrase, a sentence, an entire section, or over the whole piece of music. Of course, in these cases there are ele-

ments that are being left out. I call these cases filtered Farey sequences. In a complete Farey sequence, a rhythmic pattern is represented by the onsets between 0 and 0.5. *And*, it continues with the retrograde version of the pattern between 0.5 and 1. Of course, the *entire* Farey sequence, when read backwards, would then produce the same rhythm as the forward version. Olivier Messiaen called this type of symmetric rhythm *non-retrogradable*[1] (Messiaen 1995). Many examples of this technique can be found throughout his works, for example in the fifth movement of his composition *Visions de l'Amen* (Messiaen 1995, p. 26). The non-retrogradable rhythms use a central duration, like a hinge, which is common to both halves: `+II- I~- -II+`, with `I` = , in pulses $\{2, 2, 3, 5, 3, 2, 2\}_{19}$. To model these kinds of rhythms one needs to filter a Farey sequence in such a way that it keeps a mirror symmetry.

For the notation of any rhythm as a sequence of durations that is measured in pulses, one would set all fractions in F_n to their lowest common denominator. Then, the sequence of their numerators denotes the relative duration of every note of that rhythm. For its use in audio synthesis, one could simply multiply each numerator with the period assigned to the shortest duration. In order to model an non-isochronous meter for example, the period of pulsation of the N-cycle has to be at least around 0.1 s long according to London (2012, p. 27), see also Sect. 8.3.2, p. 121.

Another question is how many distinct note onsets occur in a polyrhythm that is modelled by F_n. The answer is, of course, given by the length of F_n. It depends on Euler's totient function $\varphi(m)$, which gives the number of integers $<= m$ that are coprime to m:

$$|F_n| = 1 + \sum_{m=1}^{n} \varphi(m) \tag{7.3}$$

The totient function is calculated by using the following product of primes that are dividing n:

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right) \tag{7.4}$$

Table 7.1 shows the development of $|F_n|$ for the first few n. Note that if n is a prime number, $n - 1$ fractions are inserted into F_{n-1}, the previous Farey Sequence.

Rhythms proceed in sequences of notes with specific durations. The inter-onset interval (IOI) descibes the link between them. A meter consists either of isochronous or non-isochronous beat cycles. Meters contain beat cycles with subdivisions. Similarly, one can link multiple Farey sequences together to form a longer sequence. And, by doing so, one would make sure that the value of the ratios remains between 0 and 1. The resulting sequence would not be a Farey Sequence anymore, because one or more elements will be missing. The sequencing of smaller F_ns, as described above, always creates a subset of a larger F_k. Let us consider the case of two identical Farey sequences F_n being linked together. The ratios in both sequences range

[1] In poetry a similar, non-reversable, mirror symmetric structure of letters, sounds or syllables is referred to as a *palindrome*, for example the famous "Sator Arepo Tenet Opera Rotas".

Table 7.1 Development of the length of F_n for some values of n

n	$\lvert F_n \rvert$	n	$\lvert F_n \rvert$	n	$\lvert F_n \rvert$
1	2	11	43	30	279
2	3	12	47	40	491
3	5	13	59	50	775
4	7	14	65	60	1103
5	11	15	73	70	1495
6	13	16	81	80	1967
7	19	17	97	90	2481
8	23	18	103	100	3045
9	29	19	121	200	12233
10	33	20	129		

between 0 and 1. In order to form the subset of a larger Farey sequence, F_{n+n}, the ranges need to be scaled. The first sequence will have a range between 0 and 0.5, the second sequence will have a range between 0.5 and 1. Using two instances of F_4 as an example, let us transform them into a subset of F_8: First, the two sets of fractions have to be scaled by a factor of two, and the resulting fractions reduced to their lowest terms. We introduce the notation $F_n[x, y]$ with $x, y \in \mathbb{R}$ to indicate the scaling of the fractions in F_n into the range given by the two real numbers x, y.

$$\{F_4, F_4\} = \left\{ \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1} \right\}, \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1} \right\} \right\}$$

$$\{F_4[0, 0.5], F_4[0.5, 1]\} = \left\{ \left\{ \frac{0}{1}, \frac{1}{8}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{3}{8}, \frac{1}{2} \right\}, \left\{ \frac{1}{2}, \frac{5}{8}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{7}{8}, \frac{1}{1} \right\} \right\}$$

Then, by process of unification, the duplicate term $\frac{1}{2}$ is deleted, resulting in an ordered subset of F_8:

$$F_4[0, 0.5] \cup F_4[0.5, 1] = \left\{ \frac{0}{1}, \frac{1}{8}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{7}{8}, \frac{1}{1} \right\} \subset F_8$$

For comparison, here is F_8 :

$$F_8 = \left\{ \frac{0}{1}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{1}{1} \right\}$$

It is evident that the onset times of the initial sequences change, but the ratios of the durations (IOIs) remain the same due to the linear scaling. Forming a subset like the above is equivalent to taking a higher order Farey Sequence and filtering out

those fractions, which have denominators composed of higher prime numbers, 5 and 7 in the above example. Of course, the principle of composing longer sequences also works when the order n of the Farey sequences is not equal. We note that a subset of F_n, which is obtained by sequencing, is a filtered Farey sequence, where the filtering method will involve the prime number composition of the denominators of the fractions in F_n. The filtering and the concatenation of Farey sequences can be used both in order to represent musical rhythm and meter. We found that a filtered Farey sequence can contain linearly scaled versions of complete Farey sequences of lower order. This is possible because the Farey sequence has a self-similar structure. Finally, if F_n can model musical meter and rhythms within a single bar, then sequenced instances of F_n are able to model an entire score.

7.2.1 *Building Consecutive Ratios Anywhere in Farey Sequences*

Given that the distance between two consecutive elements of a Farey sequence is always $\frac{c}{d} - \frac{a}{b} = \frac{1}{db}$ and if we only know the fraction a/b given in its lowest terms, then building the harmonic series:

$$\{\ldots, -\tfrac{1}{3b}, -\tfrac{1}{2b}, -\tfrac{1}{b}, \tfrac{1}{b}, \tfrac{1}{2b}, \tfrac{1}{3b}, \ldots\}$$

gives all possible distances $|\frac{1}{kb}|$ from a/b with $k \in \mathbb{N}$ where all fractions adjacent to $a/b \in F_n$ could be found, no matter how high n becomes. Of course, not all coefficients k lead to adjacent fractions. The search for the fraction c/d adjacent to a/b has to be constrained by the relation: $ad - bc = 1$ if $c/d < a/b$. Similarly, $bc - ad = 1$ if $c/d > a/b$. If we furthermore choose k so that d becomes the largest value $<= n$ fulfilling that relationship, then we always find the preceding and the succeeding ratio adjacent to $a/b \in F_n$. This has already been proven by Hardy and Wright (2008), p. 25. We have therefore implemented an algorithm that is able to construct subsets of F_n of any length $1 <= m <= |F_n|$ given only one member of F_n as a seed value. This is useful because for large n the construction of the *entire* list of ratios $\in F_n$ can take a considerable amount of computing time and memory space, because the length of F_n grows rapidly with n, see Eq. 7.3 and Table 7.1.

7.2.2 *The Farey Sequence, Arnol'd Tongues and the Stern–Brocot Tree*

The Farey sequence is a subtree of the Stern–Brocot tree (Graham et al. 1994; Calkin and Wilf 2000; Sloane 2011). The Stern–Brocot tree is generated by starting with the fractions $0/1$ and $1/0$. At each iteration, one inserts a mediant fraction between each pair of fractions obtained from the previous iteration. This generating procedure

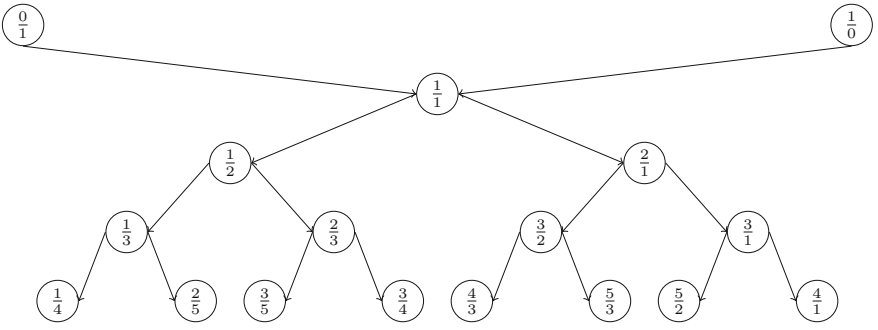


Fig. 7.2 The Stern–Brocot tree. Its left-hand branch growing from $0/1$ and $1/1$ is called the Farey tree

is different from the Farey sequence, because the size of the denominator is not constrained as in Eq. 7.1.

Starting with the root $1/1$ and per subsequent iteration, each new mediant fraction forms a new node, which is then connected to the nearest fraction or node obtained from the previous iteration. This builds the following binary tree, see Fig. 7.2. The left-hand side of the tree, going down from $1/2$, has fractions, which are elements of Farey Sequences. But, on the same level of the tree there are fractions which may belong to two (or more) different Farey sequences, for example the fourth iteration adds the fractions $\frac{1}{4}$, $\frac{2}{5}$, $\frac{3}{5}$, $\frac{3}{4}$. Although $\frac{1}{4}$ and $\frac{3}{4}$ complete Farey sequence F_4, $\frac{2}{5}$ and $\frac{3}{5}$ are only two of the five fractions that would be needed to obtain F_5. We see that the nth iteration for the Stern–Brocot tree does not generate the Farey sequence F_n on the left-hand branch of the tree. It only generates a subset. The literature refers to this branch as *Farey tree* (Large and Kolen 1995; Schroeder 1991, p. 336). It says that the integer ratios of a Farey tree mark the strength and ordering of mode locking regions, also known as Arnol'd Tongues, in a dynamic system of weakly coupled oscillators.

7.2.3 Farey Sequences and Musical Rhythms

We have seen that it is possible to use the Farey sequence as a form of representation of rhythms in Western notation. We will show in Chap. 9 that it can also be used in order to notate onsets recorded from music performances. The method uses Farey sequences to form grid points for rhythm quantisation; the outcome of the quantisation process is then a sub-set of some Farey Sequence F_n. We showed already in Boenn (2007) that Farey sequences can represent rhythms of many diverse styles from different historical and cultural contexts, because of their scalability and self-similarity and because appropriate filters can be found. Results obtained by Desain and Honing (2003) and Papadelis and Papanikolaou (2004) suggest that

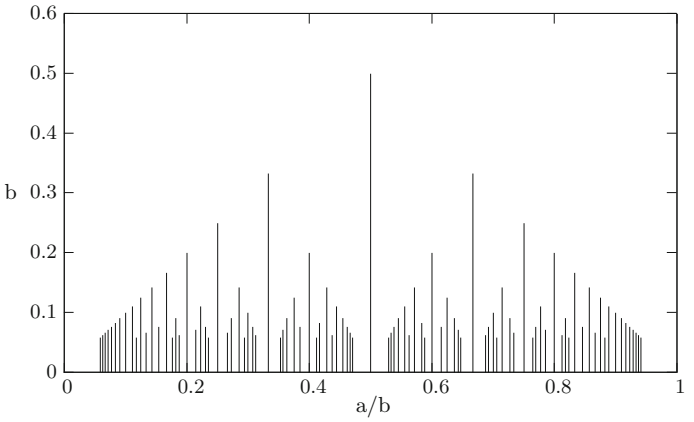


Fig. 7.3 Correlation of $a/b \in F_{17}$ and $1/b$ in the interval $]0 \ldots 1[$

human listeners categorise performed rhythms in such a way that classes of short rhythmic patterns are formed in the vicinity of relatively small integer ratios. Filtered Farey Sequences are capable of representing such categories.

Graham et al. (1994) have shown that the Stern-Brocot tree can serve to approximate decimal fractions by integer fractions, with arbitrary precision. See also (Martelli et al. 2005, p. 675) for a Python algorithm that approximates floating point numbers into Farey fractions. One of the main advantages of the Farey sequence is that it is scalable to different musical and cognitive timescales and does not rely on context-related manifestations of meter and bar. However, metrical hierarchies on any level can be modelled easily by using filtered Farey sequences.

Musically Relevant Structure of the Farey Sequence

Figure 7.3 shows a plot of F_{17} that correlates the position of the ratios a/b in the interval $]0 \ldots 1[$ with the unit fraction $1/b$ built from their corresponding denominators. The unit fractions represent the metrical subdivision by b of the reference duration 1.

The plot nicely illustrates the self-similarity of a Farey sequence. The visualization also shows that the smaller the denominator, the larger are the symmetric gaps around the x-position of the ratio, i.e., the smaller b of the ratio a/b, the greater the distance to the next ratio with the denominator $b+1$. We believe that it is due to those relatively large gaps surrounding small integer ratios that they form a zone of higher probability for rhythm and meter perception to lock onto them (mode locking). Onset time ratios that fall into these zones are shifted towards the ratio that exhibits a local peak of attentional energy (Large and Palmer 2002). Another way of looking at the space immediately around small integer ratios is to note that the denominators in the immediate vicinity of those ratios have the tendency to become larger and larger, in fact when the Farey index n goes to infinity, then the denominators of the ratios, who mark their positions on the x-axis close to small integer ratios, also go to infinity.

One can see this phenomenon indicated in Fig. 7.3. Around the value $x = 0.5 = \frac{1}{2}$ there is a tendency for denominators of surrounding fractions to become gradually larger and larger as one approaches $x = 0.5$ using fractions $\frac{a}{b}$ close to $\frac{1}{2}$, i.e., the value of $\frac{1}{b}$ becomes smaller and smaller. Can it be that this value of $\frac{1}{b}$ at normalized time positions given by $\frac{a}{b}$ is perceptually relevant? There is experimental evidence that zones of higher attention around small integer ratios indeed exist (Large and Palmer 2002). This will be an important result to consider in the design of software for quantization and tempo tracking of musically performed events.

Composers who want to "stay away" from those simple ratios will need to leave a considerable amount of space around these zones of attraction, or they would need to blur the emergence of small integer ratios by using a dust of events around time points of high metrical expectancy. This technique can be seen in many works of 20th century classical music, for example in the scores of Ferneyhough, Nono, or Ligeti, who use higher degrees of metrical subdivision on a regular basis, and in nested combinations. One can also expect that the onset times from musical performances of rhythms that are aligned to a metrical grid exhibit characteristic deviations from the small integer ratios used for notation. Above we have seen that even the smallest deviations produce ratios with relatively high denominators and with relatively high prime numbers as their components. We identify this as the main problem for rhythm transcription from performance data. The inter-onset ratios recorded from performances have to perform a small shift towards a ratio that can be used for notation. Usually, such ratios only involve small prime number divisors of their denominators, so-called k-smooth numbers (Berndt 1994; Blecksmith et al. 1998).

To continue our analysis, there are various *accelerandi* and *ritardandi* encoded in every F_n, for example there are patterns that form visible triangles between larger reciprocals and smaller ones in their surrounding area, see Fig. 7.3. These patterns are formed by monotonically increasing or decreasing values of the denominators. The increasing or decreasing tendencies can overlap each other. The gaps between the ratios forming those triangles are always on a logarithmic scale, hence the impression of accelerated or reduced tempo that becomes evident through the sonification of these patterns. These structures are mirrored around $1/2$ in every F_n with $n > 1$.

If we take the space between 0 and 1 as the normalized duration of a musical sequence with a period of x milliseconds, then only a certain degree of subdivision is perceptually relevant as an element of a rhythmic figure. If we go a step further and investigate also compound rhythms that emerge between various levels of subdivisions then we must take into account the subdivision of the beat based on the least common denominator between those two levels of subdivision of the beat that are involved in producing the shortest inter-onset intervals. The subdivisions that produce the shortest IOI between their elements are easy to find. For any F_n, the shortest IOI will be given by $\frac{1}{n} - \frac{1}{n-1}$. In order to apply a perceptual constraint, one has to filter out certain subdivisions so that IOIs do not exceed a certain time period x. Normally, 0.1 s are considered to be the perceptual threshold (London 2012). We will give more details about this threshold in the next chapter.

7.3 Filtered Farey Sequences

7.3.1 Introduction

A Farey sequence can represent the onset times and relative durations of metric subdivisions of a beat. The concatenation of various F_n into a sequence forms subsets of a higher order F_{n+k}. Therefore, appropriate filtering methods can model any musical rhythm as a filtered instance of F_{n+k}. Such a filtered sequence that represents onsets and IOIs of musical events can model the compound rhythm of any musical sequence that is either performed, read from a score or algorithmically generated by a computer. Normalized onset times can be converted into integer ratios via continued fraction expansion. A maximum of ten recursions within the expansion algorithm provides sufficient precision. The highest denominator in a sequence of such onsets determines the order n of the Farey sequence F_n that forms the superset for all onsets.

The compound rhythms and normalized onset times of a musical performance form a subset $x \subset F_n$. F_n will contain another subset y that holds the onset times of the score representation of the performed rhythm. We expect x to have a large amount of elements with relatively large integers as denominators, whereas the set of score onsets y features mostly fractions with relatively small denominators. The highest prime number component of all metrical subdivisions within the score indicates the smoothness of the denominators found in y. A positive integer is k-smooth if its prime divisors are $\leq k$ (Berndt 1994; Blecksmith et al. 1998). In our case we are interested in the smoothness of the denominator of a fraction in y. For example, if the highest prime number subdivision of all note durations within the score is 3, then all the denominators of the fractions that represent note onsets must belong to the *3-smooth* numbers, which means they can only be products of the powers of 2 and 3. The performed onsets in x introduce deviations from the score onsets because of expressive timing, therefore one can expect that the fractions in x will have relatively large denominators with prime divisors > 3. Only a machine performance of the score onsets y would return an identical set $x = y$. Therefore it is appropriate to look at x and y as subsets of F_n with a high probability of the relation $x \neq y$. In our previous analysis we have seen that the subsets x and y are both filtered Farey sequences or scaled versions of a complete Farey sequence of lower order. The question is whether there can be an algorithm that can analyze the sets x and y in such a way that it is able to map each member of x uniquely to a member in y. This would allow us to use a Farey sequence not only to represent the onset times of an entire score, but also to find a mapping of one subset of a Farey sequence that represents performed onset times to a subset of F_n, which represents the score onset times of that performance.

In the following sections, I am going to analyse music with the help of filtered Farey Sequences. Then, we will arrive at a few definitions for filtered Farey Sequences, which are useful for modelling musical rhythm and meter.

7.3.2 Polyrhythms

A very common case of polyrhythm in Western music is the so-called *hemiola*. This two-voice rhythm has a proportion of 2 : 3 and it can be modelled with F_3. This kind of hemiola can be heard very often in ornamented cadences of the Baroque era. Two ternary bars can be combined to form a 'super-ternary' structure with an emphasis of the 1st and 3rd beat of the first bar, followed by the 2nd beat in the second bar. In history, one can find many examples of this rhythmic device. From the romantic period, there is the beginning of Robert Schumann's 3rd Symphony in E-flat major. Three consecutive hemiolas are presented right at the opening of the first movement. The actual flow of 3/4 meter of the piece is only established in bar 7.

Farey sequences can represent complex polyrhythms. Table 7.2 shows a list of common two-layered polyrhythms developed for modern education (Marron 1990). It is also conceivable to extract multi-layered polyrhythms from Farey sequences. Table 7.3 shows a few examples. The number sequences represent the inter-onset times between consecutive beats of the polyrhythm. Of course, with increasing order of the Farey sequence, one observes an increase of the 'gaps' we have mentioned before: The high numbers at the beginning, end, and in the middle of the sequence are good indicators for the tendency of the fractions to only very slowly converge around $\{0/1, 1/2, 1/1\}$. It also becomes clear that the greater the prime factors are, which take part in forming the common denominator, the larger the numerical size of the inter-onset intervals (IOIs). In addition, there is a human limit of the perception of differences between the lengths of IOIs. London (2012, p. 27) quotes evidence that this limit is around 0.1 s. Even with this smallest duration of a pulse, a musical rendering of F_8 will become a slow movement, with IOIs ranging between 1.5 and 10.5 s; the entire sequence would last 84 s! If one would increase the tempo, the differences between the smaller durations would become more and more blurred,

Table 7.2 Polyrhythms and their corresponding filtered Farey sequences

Pattern	Ratio	Filtered F_n
`I:I`	3:2	F_6
`I:II:I`	6:4	F_{24}
`-.II.-`	4:3	F_{12}
`II:II`	5:2	F_{10}
`III:III`	7:2	F_{14}
`IIII:IIII`	9:2	F_{18}
`-I.-.I-`	5:3	F_{15}
`H.-II-.H`	5:4	F_{20}
`--.I-I.--`	7:3	F_{21}
`H-.HIIH.-H`	7:4	F_{28}
`HH.-HIIH-.HH`	9:4	F_{36}

Table 7.3 Some complete Farey sequences, their transcriptions and IOIs

Pattern	F_n	Inter-onset intervals (IOIs)
`I:I`	F_3	{2, 1, 1, 2}
`-.II.-`	F_4	{3, 1, 2, 2, 1, 3}
`H~H~H--~IH-~--~-HI~--H~H~H`	F_5	{12, 3, 5, 4, 6, 6, 4, 5, 3, 12}
`H~H~II-I~-H-~--~-H-~I-II~H~H`	F_6	{10, 2, 3, 5, 4, 6, 6, 4, 5, 3, 2, 10}
	F_7	{60, 10, 14, 21, 15, 20, 28, 12, 30 30, 12, 28, 20, 15, 21, 14, 10, 60}
	F_8	{105, 15, 20, 28, 42, 30, 40, 35, 21, 24, 60 60, 24, 21, 35, 40, 30, 42, 28, 20, 15, 105}

turning into the statistical perception of the events as a gesture. Many such gestures exist in music, for example, the trill, the arpeggio, the flam, the drum roll, the tremolo, etcetera. In notation, however, there is a clear distinction between subdivisions that are written out and musical gestures such as the trill, which is abbreviated as one symbol. The solution with regard to polyrhythms and Farey sequences is to apply a filter for the maximum prime number used in a Farey sequence. I will introduce such a filter in Sect. 7.3.6.

Table 7.3 shows the compound rhythm representation of the polyrhythm 6:5:4:3:2, i.e. F_6, displayed by Eq. 7.5. Each fraction represents a note onset. Compare this with Fig. 7.4, which distributes the subdivisions onto different voices.

$$F_6 = \left\{ \frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1} \right\} \tag{7.5}$$

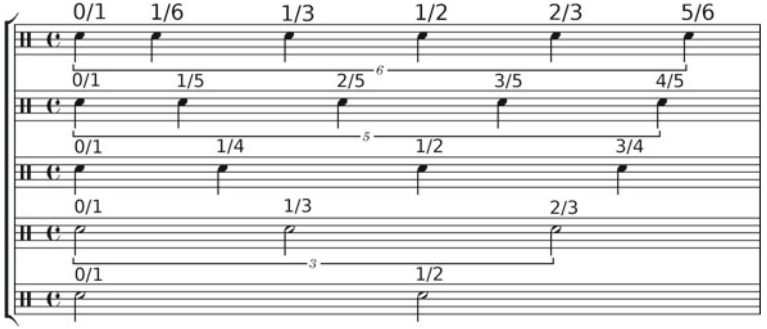


Fig. 7.4 The polyrhythm of F_6 split into 5 voices. The entire bar represents the space between $\frac{0}{1}$ and $\frac{1}{1}$. Each subdivision is represented by its own voice. The merger of all five voices into a single line is shown in the left column of Table 7.3

Ockeghem

Polyrhythmic concepts have been redeveloped by 20th century Classical composers who were influenced by non-Western musical cultures, for example György Ligeti, see Taylor (2012). He also might have been inspired by 15th century composer Johannes Ockeghem who composed his *Missa prolationum* using a sophisticated structure of two simultaneous double-canons, whereby, for each canon, two voices use the same pitch sequences and rhythmic values but each of the voices execute them using their own tempo. All voices do this simultaneously, thereby creating a complex vocal polyphony, see Tables 7.4 and 7.5.

Creation of four different tempi using the same sets of notes and rhythms was possible because the duration of the brevis, which translates into today's double whole-note, was not fixed as in Western notation today. There were many different options for the composer of that era, as for example outlined in the famous 14th

Table 7.4 Onset times of the metrical subdivisions underlying the first voice in Table 7.5 expressed as F_{36}. The structure of subdivision is ◯, *tempus perfectum cum prolatione imperfecta*. This meta-cycle encompasses 6 mensurations of 3 × 2 half notes at the beginning of Johannes Ockeghem's Credo. The beat period is $\frac{1}{36}$

I	$\frac{0}{1}$						$\frac{1}{6}$						$\frac{1}{3}$				
II			$\frac{1}{18}$		$\frac{1}{9}$				$\frac{2}{9}$		$\frac{5}{18}$				$\frac{7}{18}$		$\frac{4}{9}$	
III		$\frac{1}{36}$		$\frac{1}{12}$		$\frac{5}{36}$		$\frac{7}{36}$		$\frac{1}{4}$		$\frac{11}{36}$		$\frac{13}{36}$		$\frac{5}{12}$		$\frac{17}{36}$
I	$\frac{1}{2}$						$\frac{2}{3}$						$\frac{5}{6}$					
II			$\frac{5}{9}$		$\frac{11}{18}$				$\frac{13}{18}$		$\frac{7}{9}$				$\frac{8}{9}$		$\frac{17}{18}$	
III		$\frac{19}{36}$		$\frac{7}{12}$		$\frac{23}{36}$		$\frac{25}{36}$		$\frac{3}{4}$		$\frac{29}{36}$		$\frac{31}{36}$		$\frac{11}{12}$		$\frac{35}{36}$

Table 7.5 Polymetric structure of beats from Ockeghem's Missa prolationem creating a meta-cycle of 36 beats. '•' marks the beginning of a new mensuration (equivalent to today's downbeat), '○' marks a 2nd level beat, '·' marks a 3rd level beat

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Discant	•	·	○	·	○	·	•	·	○	·	○	·	•	·	○	·	○	·
Alto	•	·	○	·	•	·	○	·	•	·	○	·	•	·	○	·	•	·
Tenor	•	·	·	○	·	·	○	·	·	•	·	·	○	·	·	○	·	·
Bass	•	·	·	○	·	·	•	·	·	○	·	·	•	·	·	○	·	·
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
D	•	·	○	·	○	·	•	·	○	·	○	·	•	·	○	·	○	·
A	○	·	•	·	○	·	•	·	○	·	•	·	○	·	•	·	○	·
T	•	·	·	○	·	·	○	·	·	•	·	·	○	·	·	○	·	·
B	•	·	·	○	·	·	•	·	·	○	·	·	•	·	·	○	·	·

century treatise 'Ars Nova' by Philippe de Vitry (1964). The brevis can be interpreted either as 3 or 2 semibreve, and those further into 3 and 2 subdivisions called half notes, so we can subdivide the brevis into 3×2, 2×2, 3×3, or 2×3 half notes. This was indicated in the vocal score with different mensuration signs, from which today's time-signatures developed.

Ockeghem uses all four possibilities simultaneously, thus introducing a proportion of 6:4:9:6 between the voices. The voices need 36 half notes in today's notation to complete the full cycle of patterns and start on the same 'downbeat' again (de la Motte 1981). The complete cycle translates into a *filtered* Farey sequence F_{36}. Given that there were no complete scores written in the modern sense, only the vocal parts, it demonstrates the ingenuity of Ockeghem's mastery of polyrhythms and how the application of different meters on the same melodic material affects the metric feel of the individual voices. Bear in mind that, in the Renaissance, the meter does not affect the smaller note values below the half note (minim). Therefore, the durations of the quarter notes (*Semiminim*) and of the eighth notes (*fusa*) are the same in all four voices, regardless of the meter.

The hierarchical pattern of subdivisions in Table 7.4 shows some interesting properties that demonstrate a general method for the construction of such patterns. Due to space constraints we can only present the *Discantus*, the highest voice in Ockeghem's composition, but the principle can be easily adapted for the other voices. Starting with the highest metrical level (I) in Table 7.4, which has a period of 6 half notes, we see that the largest denominator is equal to the number of mensurations required to finish the meta-cycle together with all the other voices. All other denominators on level I are divisors of this number. The next level (II), with a period of two half notes, shows us how many subdivisions are contained in the meta-cycle on that particular level. This is again indicated by the largest denominator of the level and all other denominators are its divisors, but *excluding* those already contained on higher levels. This principle repeats itself on all metrical levels. For each of the denominators b in F_{36}, their numerators also follow a simple principle: they traverse the ordered set of numbers coprime to b. In our example, for $b = 18$, its numerators a, in the order of their appearance in F_{36}, are $\{1, 5, 7, 11, 13, 17\}$. This is rooted in the general properties of the Farey sequence, see Eqs. 7.3 and 7.4 on page 86.

In general, rhythmic concepts that depend on a regular pulse deal with various degrees of regularity and irregularity that are applied to durations, tone colours (pitch and dynamics) and accentuation (dynamics and articulation). We distinguish further between repetitive regular patterns and repetitive irregular patterns. The former being obtained by selecting a single element out of F_n and repeating it unchanged in cycles, or by filtering F_n in such a way that only those fractions remain that represent equal inter-onset times. If this process is varied after each cycle then it changes into a pattern of irregular duration contrast. Contrasts can also be achieved by using the other musical parameters: accentuation, tone-colour, chord density and dynamics.

Repetitive irregular patterns, on the other hand, are created through a series of *different* elements taken from F_n, for example by using a rule-based selection process. A repetition of this pattern may occur after each cycle. If the series changes after

each cycle then this process shows an irregular duration contrast. Of course, one can substitute duration by using a different musical parameter instead.

Stravinsky

Igor Stravinsky's music to the ballet *The Rite of Spring*, which was first performed in 1913, contains a famous polyrhythmic passage called the *Procession of the Sage* (Stravinsky 1967, p. 62). Figure 7.5 shows six layers of pulsations that are associated with different melodic and harmonic material. These are being played by different instrumental groups.

The proportions of these pulsations are 12:8:6:4:3 in relation to the length of one bar in 6/4 meter. When looking at the layering of the melodic material one detects even longer periods for superimposed melodic patterns, which are repeated over and over again in a so-called *ostinato*.

The climax of the *Procession of the Sage* is a 48 quarter notes long passage at rehearsal number 70 (Stravinsky 1967, p. 63), where all pulsations start to sound together. They complete one full cycle of six 6/4 bars and then end abruptly with a general silence. Table 7.6 shows the lengths of the different periods of pulsations and cyclic patterns that Stravinsky employs over six complete bars in a 6/4 meter.

From this analysis we can derive how a Farey sequence can model the rhythmic structure of the polyrhythms in the *Procession of the Sage*. The compound rhythm in Fig. 7.5 can be represented by F_{24}, which is filtered in such a way that only the subdivisions by 24, 12, 8, 6, 4 and 3 prevail. The subdivision by 24 ensures the inclusion of the sixteenth notes of the Oboe pattern. However, due to the participation of the subdivision by 8 in the Guero pattern, only two sixteenth notes are not covered by any of the other pulse onsets. They have the onset values $11/24$ and $19/24$. The

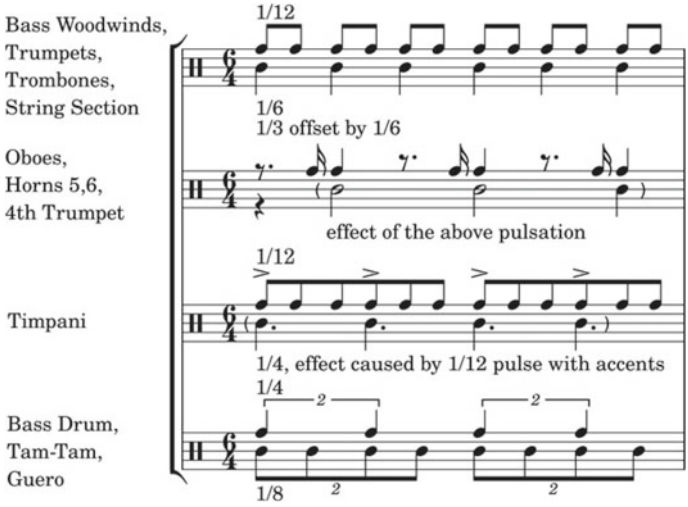


Fig. 7.5 Polyrhythmic passage in the The Rite of Spring at rehearsal number 70. Pulsation lengths are given in integer fractions relative to the length of one bar

filtering of F_{24} can thus be refined to allow only the subdivisions by 12, 8, 6, 4 and 3 with subsequent addition of $11/24$ and $19/24$; note that both numerators are prime. All denominators in the filtered Farey sequence are 3-smooth. From the series of 3-smooth numbers up to 12, which is $\{1, 2, 3, 4, 6, 8, 9, 12\}$, only the subdivision by 9 is missing, therefore the only power of three that is allowed to participate is 3^1. In musical terms, the subdivision by 9 would have been a pulsation in quarter note triplets which would have coincided with the normal quarter notes at onset times $\{0/1, 1/3, 2/3\}$. Perhaps this would have created an unwanted metrical emphasis at those ternary points of the meter 6/4. It shows that for the subdivisions involved in musical rhythms, not only the prime number factors are important, but also their exponents are equally important. In essence, this polyrhythm has the structure of a super-hemiola, because we find the superimposed proportions of $2/3, 4/6, 8/12$. Interestingly, all voices participate in one of these hemiolas, and $1/2$ is the only point, apart from the downbeat, where *all* pulses from each voice must coincide.

In Stravinsky's second recording for CBS of this section, in which he conducted himself, the period for a quarter note is approx. 0.386 s, or roughly 155 BPM (Stravinsky 1960). Based on this tempo, we can analyze how the perceptual timescales map to the voices of the polyrhythm. The tempo-metrical grid in Table 7.6 shows the analysis.

Table 7.6 Overview of proportions used in the *Procession of the Sage*. Ratios are related to the length of 6 bars in 6/4, or a total of 48 quarter notes

Notation	Context	Proportion	Period (s)	Tempo
`[.]`	Part of off-beat pattern of Oboes and Brass	1/192	0.0965	
`.`	Pulsation in lower woodwinds and strings	1/96	0.193	310.88 BPM
`[-]`	Guero only	1/64	0.2895	207.25 BPM
`I`	Violin I and combination of Trumpets and Trombones	1/48	0.386	155.44 BPM
`-`	Bass drum, Tam-Tam and Timpani	1/32	0.579	103.63 BPM
`H`	Off-beat pulsation in Oboes and Brass	1/24	0.772	77.72 BPM
`H~I`	Main note of the tuba theme	1/16	1.158	51.81 BPM
`W`	Phrase periods in most melodic patterns, e.g. of the tubas	1/12	1.544	38.86 BPM
`W~W`	Periods of the horns, Trumpets and violin patterns	1/6	3.088	
`W~W~W~W`	Long period of the tuba theme	1/3	6.176	

The extraordinary result is that Stravinsky uses the entire bandwidth of musical timing structures ranging from roughly 0.1 s for the smallest subdivision up to slightly above 6 s for the largest ternary bar structures (see Sect. 8.3 and especially Fig. 8.5 on p. 123). At the centre we find one period that is exactly in the range of the maximum pulse salience (0.597 s), which is the pulsation carried out by the combination of Bass Drum, Tam-Tam and Timpani. This is a pulsation that goes against the grain of the melodic information in the violins, trumpets, horns and tubas, which is relying on the quarter note subdivision (0.386 s). Both streams together participate in one of the hemiola structures (6:4 in relation to one bar of 6/4 meter) that we have discovered earlier.

African Drumming

Sima Arom's research (Arom 1991) has been influential on some contemporary Western composers. Very successful recording and transcription methods form the basis for his analysis of African polyrhythm. We would like to translate some of the principles described in his book into the realm of Computer Music for creative purposes but also in order to demonstrate the general use of our Farey Sequence model. We are referring to Arom's classification of rhythmic structures (Arom 1991, p. 203) that are part of the African drumming traditions he investigates, but which also occur in some of our Western traditions, for example in the music by Steve Reich. We would like to present the following 'recipes' in order to create rhythms based on these principles.

Type a *Identical durations with regular accentuation*: (1) choose from F_n a series of Ratios with equal distance or delta time d. (2) choose an integer multiple of d for accents and create tables encoding steps (1) and (2).

Type b *Identical durations with irregular accentuation*: (1) choose from F_n a series of Ratios with equal delta time d. (2) choose an integer multiple i of d for accents. Choose i from a table of small integers. (3) repeat step (2) after i beats, thereby creating an irregular accentuation contrast.

Type c *Identical durations with no accentuation but regular alternation of tone colour*: (1) choose from F_n a series of Ratios with equal delta time d and decide on the series' length. (2) according to the F_n frame in (1), map an equally sized list of tone colours, where repetitions may also occur.

Type d *Identical durations with no accentuation but irregular alternation of tone colour*: (1) choose from F_n a series of Ratios with equal delta time d and decide on the series' *maximum* length in beats. (2) according to the F_n frame in (1), map lists of tone colours not larger than the series in (1), repetitions of tones are allowed.

Type e *Differing durations with regular accentuation*: (1) choose a F_n and filter to obtain desired rhythms, for example via smooth numbers. (2) Place accents within this series at equal delta times.

Type f *Differing durations with irregular accentuation*: (1) choose a F_n and filter to obtain desired rhythms, for example via smooth numbers. (2) Place accents within this series at non-equal delta times.

Type g *Differing durations with no accentuation but regular alternation of tone colour*: (1) choose a F_n and filter in such a way that a large subdivision, e.g., 1:2 always occurs. (2) map tone colours in such a way that a change in tone colours only occurs at the boundaries of those large subdivisions.

Type h *Differing durations with no accentuation but irregular alternation of tone colour*: (1) choose a F_n and filter to obtain desired rhythms, for example via smooth numbers. (2) map tone colours in such a way that a change in tone colours occurs at non-regular delta-times.

Type i *Differing durations with neither accentuation nor change of tone colour*: (1) choose a F_n and filter to obtain desired rhythms, for example via smooth numbers or probabilistic filters.

Mix : Carry out types (a) to (i) in succession and in polyphony, also apply and mix those methods with regard to other musical parameters.

7.3.3 Rhythm Transformations

We have touched on growth and decline as principles of musical form in Chaps. 2 and 4. In addition, there are sheer endless possibilities to change a rhythmic motive. Here are some general principles of variation in the following Fig. 7.6. We also map the onsets of rhythmic events to elements of the Farey sequence representing the entire length of the motive. Note that only the operations *Ellipsis*, *Augmentation* and *Diminution* preserve the initially given proportions between note durations. Even more complex shifts of proportions will arise from the simultaneous combination of selected transformations leading to a rich palette of variations. Simple yet musically powerful rhythmic motives, which can serve as the basic cells for these kinds of variations, could be taken from the repertoire of Greek verse rhythms, for example (see Table 6.1).

7.3.4 Greek Verse Rhythms

We have already encountered Greek verse rhythms in Chaps. 3 and 6. Here is now another method for creating Greek verse rhythms; one that uses filtered Farey sequences. In order to build a Iamb or a Trochee, we generalize the fact that the second note is shifted away from the centre (0.5) of the time frame (0...1), either towards the first note (0) or towards the end of the time frame (1). The former will produce a Iamb, the latter creates a Trochee. Similarly, in order to create an Anapest we would first create a Iamb, then subdivide the emerging frame between the first and second note of the Iamb. A Dactyl is created by subdividing a Trochee's frame between its second note and the end of is time frame. Note, that the subdivision again may produce a Iamb or a Trochee instead of a clean $1:2$ ratio. An Amphibrach is created by the merger of a Iamb and a Trochee. Many more rhythms can be created

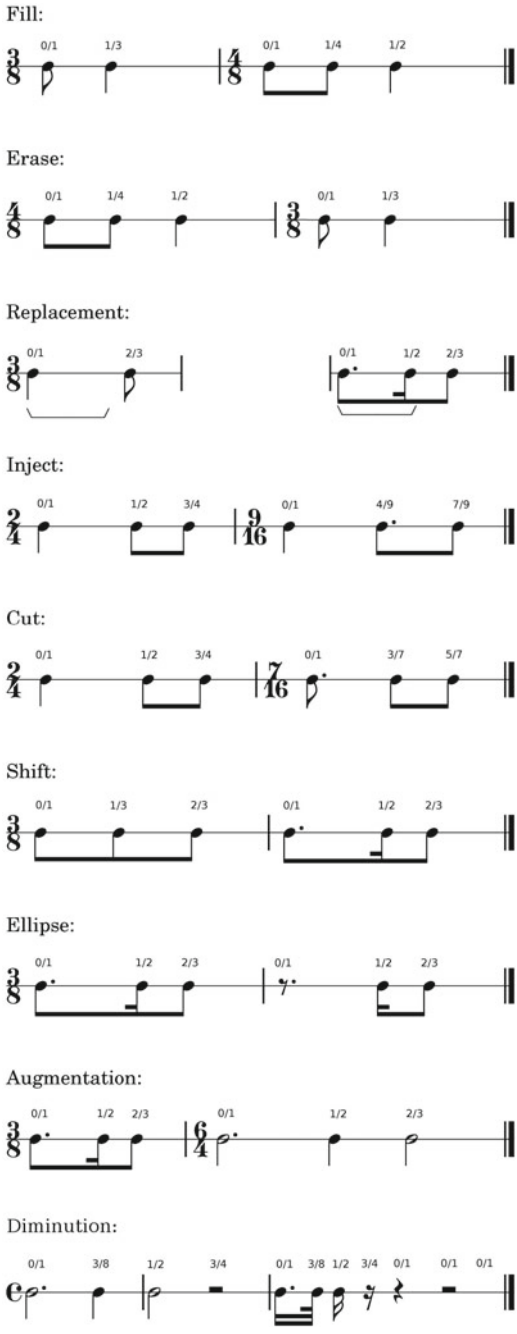


Fig. 7.6 List of possible operations in order to create rhythmic variations of small rhythmic motives. The original rhythm is displayed on the left-hand side, its transformation is shown on the right-hand side. Onsets are marked by elements of a Farey sequence whose length equals the total duration of the rhythmic cell

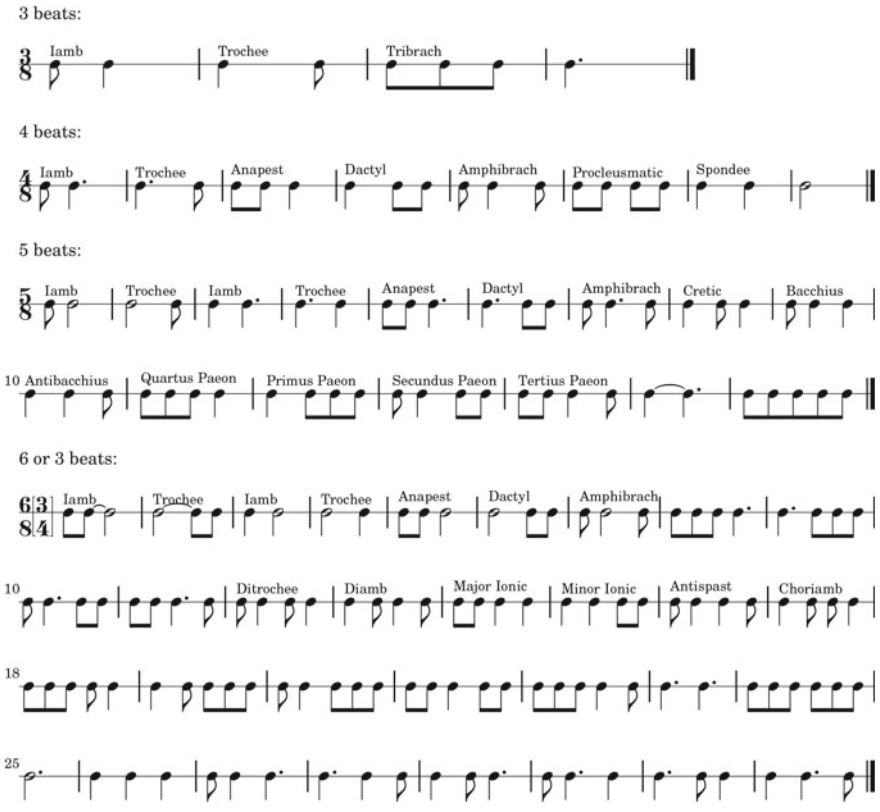



Fig. 7.7 Systematic chart of all Greek verse rhythms. The chart can continue infinitely as the number of beats per bar grows with higher $n \in \mathbb{N}$

in a similar way, all of them can be represented by F_n. For example, if one wants to work with rhythms that are based on short-long patterns that match precisely a grid of n beats, then, given a number of n beats per bar, the number of possible distinct patterns is 2^{n-1}.

The following Fig. 7.7 represents our systematic approach for the generation of all Greek verse rhythms and their variants. The new aspect of our representation is that, for example, the Iamb rhythm can manifest itself in different proportions.

When we look at all 4-beat rhythms, where 'short' is defined as an eighth note $\rightarrow a$ and 'long' is a quarter note $\rightarrow b$, then the 4-beat rhythms can be built entirely from the sets ab, aab, $aaaa$ and bb, as illustrated in Fig. 7.7.

Similar tables for all 3-beat, 5-beat and 6-beat rhythms, which are formed according to long-short patterns, lead to the conclusion that they can be seen in a more generalized view as filtered Farey sequences. The ratios that are drawn from a sequence F_n are all multiples of the ratio $\frac{1}{n}$ in $[0, \ldots, 1[$ excluding 1. For example, within F_4 we find the list for the 4-beat patterns: $\{\frac{0}{1}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$, see Eq. 7.6.

$$F_4 = \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1} \right\} \tag{7.6}$$

It appears that, based on a list of ratios drawn from F_n, all rhythmic patterns, which belong to the n-beat group are generated by a combination of those ratios, where $\frac{0}{1}$ is always picked as the first element of the beat-pattern. One can identify several aspects of this combinatorial problem. First, the number of possible combinations C_n that form all patterns of n beats can be written in the form

$$C_n = \sum_{k=0}^{n-1} \binom{n-1}{k},$$

which is equal to the sum of the $(n-1)^{th}$ line of the Pascal triangle. Looking at the series of binominal coefficients, they clearly indicate how many patterns belong to a group that contains an equal number of elements. For example, the 5-beats table contains $\sum_{k=0}^{4} \binom{4}{k} = 16$ different patterns that belong to 5 different groups of patterns. We have therefore $\binom{4}{0} = 1$ pattern of the set c, $\binom{4}{1} = 4$ patterns of the set ab, $\binom{4}{2} = 6$ patterns of the sets aab and its complement abb, $\binom{4}{3} = 4$ patterns of the set $aaab$ and $\binom{4}{4} = 1$ pattern of the set $aaaaa$. If one looks at the indices of the ratios, starting with $n = 0$ as the first element, it becomes clear that there is a way to construct all patterns algorithmically. One can ignore the 0th index because the ratio $\frac{0}{1}$ is always the first element in a pattern. Then, it is possible to take the remaining indices $S_n = \{1, \ldots, n\}$, and to perform all combinations in lexicographical order by using all possible subsets of S_n. Finally, the ratio $\frac{0}{1}$ is prepended to each one of those permutations.

7.3.5 *Filters Based on Sequences of Natural Integers*

A musical meter can be described in terms of the distances between peaks of attendance with a focus on the beat level. It is expressed as a relatively small integer multiplied by some constant value for the period of the of smallest pulsation, i.e., what London (2012) calls the N-level. For example, the sequence {3, 3, 2, 2, 2} describes a meter with non-isochronous beats on the backdrop of a constant pulsation with the period $1/12$ and $N = 12$. Musical rhythms that emerge from an underlying metrical grid show the same property. In fact, any sequence of natural integers, including those that have been generated algorithmically, can be easily translated into a filtered Farey Sequence, which serves as the model structure for the rhythm.

The method is as follows: From the sequence of natural integers we form an accumulative set, where for each element we take the sum of the previous elements in the original sequence starting with 0, for example: {3, 3, 2, 2, 2} becomes {0, 3, 6, 8, 10, 12}. Its last integer, 12, indicates the order of F_n, so that the filtered Farey sequence becomes:

$$F_{12} \supset \left\{ \frac{0}{12}, \frac{3}{12}, \frac{6}{12}, \frac{8}{12}, \frac{10}{12}, \frac{12}{12} \right\},$$

and after reducing the terms:

$$F_{12} \supset F_6 \supset \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, \frac{1}{1} \right\},$$

This method can be applied to sequences of integers of arbitrary length. Its outcome is a filtered Farey sequence that one can always translate into notation.

7.3.6 Filters Based on the Prime Number Composition of an Integer

We have already mentioned that elements of a Farey sequence could be filtered out on the basis of the prime number composition of their denominator. A filter that would allow only k-smooth numbers in the denominator can model the normalized onset times of notes written in notation, where the highest metrical subdivision used in the score equals k. However, as we have seen in the analysis of the *Procession of the Sage*, it is important to take the prime numbers *and* their exponents into account when developing filters for Farey sequences.

Clarence Barlow developed a weighting function for the analysis of the prime-number composition of a natural integer. He invented two algorithms that measure the suitability of certain integer ratios to be used in the design of musical tuning systems: He called them the function of "indigestibility" of a natural number and a function for the "harmonicity" of an integer ratio. I have already introduced the "indigestibility" for the measurement of metric complexity in Sect. 4.5. These functions play a major role in Barlow's algorithmic composition program *Autobusk*, prominently used for works like *Çoğliuautobüsişletmesi* and others (Barlow 1984). Their purpose is to facilitate the creation of tuning systems, of pitches and of chord selection algorithms used for audio synthesis or MIDI-based sound output where individual note frequencies can be tuned with cent value precision, e.g., a semitone in equal temperament is exactly 100 cents wide. The difficulty in developing tuning systems is to find a sequence of small integer ratios that are suitable for the subdivision of a standard interval, for example for subdividing the octave with a frequency ratio 2/1. One solution could be, for example, a series of intervals according to the just tuning system {1/1, 9/8, 5/4, 4/3, 3/2, 5/3, 15/8, 2/1} (Barlow 1984; Moore 1990). We are using both functions by Barlow but apply them to the 'tuning' of onset intervals that lead to the least complex notation of a recorded performance.

Barlow's "Indigestibility" Function

To arrive at measuring the "harmonicity" of an integer ratio, Barlow developed his measure of "indigestibility" of a positive integer, $\xi(N)$, based on the number of prime

Table 7.7 Table of the "indigestibility" of the first natural numbers

n	$\xi(n)$
1	0
2	1
3	2.66667
4	2
5	6.4
6	3.66667
7	10.2857
8	3
9	5.33333
10	7.4
11	18.1818
12	4.66667
13	22.1538
14	11.2857
15	9.06667
16	4
17	30.1176

factors involved and by taking into account both the size of each prime number and the size of its exponent, see Eq. 7.7.

$$\xi(N) = 2\sum_{r=1}^{\infty}\left\{\frac{n_r(p_r-1)^2}{p_r}\right\} \tag{7.7}$$

with $N = \prod_{r=1}^{\infty} p_r^{n_r}$, p_r is the rth prime number and n_r is its exponent in the prime number composition of N. Table 7.7 shows the development of $\xi(n)$ for the first 17 integers.

When sorted after increasing values of $\xi(N)$, the first 17 natural integers have the following order:

$$\{1, 2, 4, 3, 8, 6, 16, 12, 9, 5, 10, 15, 7, 14, 17\}$$

We use this measure to determine a weight C of an integer fraction in the context of inter-onset-intervals (IOIs). The "indigestibility" of the denominator determines the metrical relevance of the fraction representing the onset of a musical event and the same can be done for the IOIs. For any given ratio a/b with $a, b \in \mathbb{N}$ we calculate:

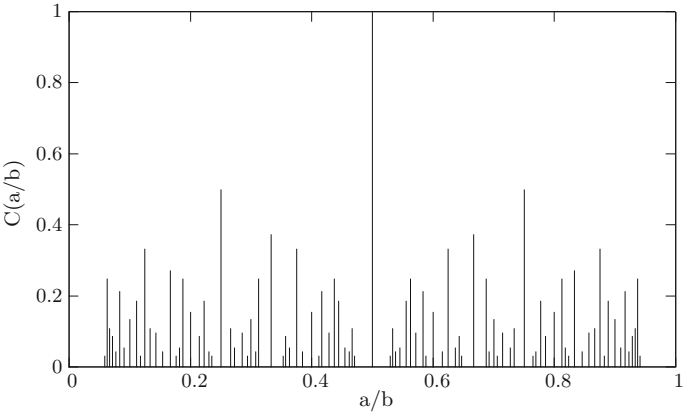


Fig. 7.8 Correlation of $a/b \in F_{17}$ and $C(a/b)$ in the interval $]0 \ldots 1[$

$$C(a/b) = \frac{1}{\xi(b)} \tag{7.8}$$

with $\xi(b)$ defined in Eq. 7.7. Figure 7.8 shows the correlation between ratios $a/b \in F_{17}$ and their weighted metrical relevance $C(a/b)$ according to Eq. 7.8.

Barlow's Harmonicity Function

Barlow uses his formula of "indigestibility" to measure the harmonicity $H(p/q)$ of a pitch interval p/q according to:

$$H(p/q) = \frac{sgn(\xi(q) - \xi(p))}{\xi(p) + \xi(q)} \tag{7.9}$$

Barlow uses the signum function in order to take the root position of the interval into account. For example, in terms of musical harmony, the fundamental of the perfect fourth, $H(3/4) = -0.214286$ is the upper note, whereas in the perfect fifth, $H(2/3) = 0.272727$, it is the lower note. The higher $|H(p/q)|$, the higher is the subjective stability of the interval (Hajdu 1993). Therefore a tritone tuned as 45/64 has only a harmonicity of 0.056391. Because our research focuses on the transcription of IOIs the polarity of those ratios are not interesting for us, therefore we simplified the Eq. 7.9 so that it always returns a positive number:

$$H(p/q) = \frac{1}{\xi(p) + \xi(q)} \tag{7.10}$$

with $H(p/q) = 1$, if $p = q$. The harmonicity measure is useful in order to evaluate a set of quantised IOIs in preparation for their transcription into notation. The problem of inferring a beat structure underlying a set of durations, based only on knowing the

durations, could be solved with the help of the harmonicity function in Eq. 7.10. We will explain the details of this application in Sect. 9.6.1 on p. 167.

7.3.7 *Metrical Filters*

A filtered F_n can represent an entire bar structure from isochronous or non-isochronous groups of beats down to all possible levels of subdivision and beyond towards meta-bar structures. Therefore, a filtered Farey sequence can be used to create the dot notation commonly used within music analysis to denote the various layers of metrical structures (Lerdahl and Jackendoff 1996; London 2012), see also Fig. 8.1 on p. 120.

Calculation of the Metrical Tempo Grid

London (2012) found interesting links between psychological thresholds of time perception in music listening and the prevalence of certain metrical subdivisions and bar structures in Western music, see also Sect. 8.3. He uses a systematic tree structure to visualize beat subdivisions in multiples of 2 and 3s and meters that also use groups of beats into binary or ternary units. The factors of subdivision and metric grouping are taken from the sequence of 3-smooth numbers. Their periods are based on the main beat period, which is derived from the tempo. Due to the psychological thresholds, this tree only shows specific subdivisions and meters that lie within the thresholds. Bars that become too long and subdivisions, which are too short are filtered out from the tree. The remaining tree forms a metrical tempo grid, see Table 7.8. The decision to use only multiples of 2 and 3s is based on the assumption that the majority of Western music to date is covered in this way. Subdivisions involving prime numbers above 5, and meters employing higher primes than 3 were indeed rarely used in Western music until approximately 100 years ago, when 20th century Classical music and Jazz started to test and push these boundaries. In order to reflect this development, one can expand London's proposed tree structure easily by using higher primes for metric subdivisions and beat groupings.

The following steps describe how to use the Farey sequence for the calculation of the metrical tempo grid. The outcome of our method is identical with the metrical tempo grid described by London (2012, p. 38).

1. Calculate F_{27}, because the highest integer for metrical subdivision is 3^3 for the majority of Western music.
2. Filter F_{27} so that only those integer ratios remain that have 3-smooth numbers as denominators.
3. From the filtered F_{27} keep only the reciprocals.
4. Duplicate every reciprocal in this set and take the inverse from the duplicated fractions.
5. Multiply every fraction in this set by a fraction that represents the duration of the beat, for example $\frac{1}{4}$.

Table 7.8 A metrical tempo grid after London with the central beat set at 650 ms or 92 BPM. This table was generated by our method using Farey Sequence F_{27}. Emphasized are all durations and meters within the range of perceptual timing limits, i.e., between 0.1 and 6 s. Periods are shown in milliseconds

Duration	Value	Period [ms]	BPM	Note name
1/108	0.00925926	24		
1/96	0.0104167	27		
1/72	0.0138889	36		
1/64	0.015625	41		
1/48	0.0208333	54		
1/36	0.0277778	72		
1/32	0.03125	81		32nd
1/24	**0.0416667**	**108**	**554**	16th triplet
1/16	**0.0625**	**163**	**369**	16th
1/12	**0.0833333**	**217**	**277**	Eighth triplet
1/8	**0.125**	**325**	**185**	Eighth
1/4 [beat]	**0.25**	**650**	**92**	Quarter
2/4	**0.5**	**1300**	**46**	Half
3/4	**0.75**	**1950**	**31**	Dotted half
4/4	**1**	**2600**	**23**	Whole
6/4	**1.5**	**3900**	**15**	Dotted whole
8/4	**2**	**5200**	**12**	Double whole
9/4	**2.25**	**5850**	**10**	
12/4	3	7800		
16/4	4	10400		
18/4	4.5	11700		
24/4	6	15600		
27/4	6.75	17550		

Table 7.8 shows, as an example, the resulting set of integer ratios representing time signatures when they are greater than $\frac{1}{4}$, otherwise they represent subdivisions of the beat. $\frac{1}{4}$ is set as the beat with a period of 650 ms. The perceptual thresholds now apply, and durations below 6 s and above 100 ms are emphasised.

Graphical Derivation of Metrical Hierarchy

It is possible to derive metrical hierarchies from a filtered Farey sequence F_n. Subdivisions of a musical duration generate an isochronous IOI, i.e., the ratio between the duration and the number of subdivisions n. Metrical hierarchies that are based on the subdivision by n emerge graphically from the tree of F_n, because integer fractions with a denominator smaller than n are rooted on a higher level of subdivision, see Figs. 7.9 and 7.10. Here, each of the higher levels would then represent a different beat cycle, and the level of subdivision by n corresponds to London's N-cycle (London 2012, p. 92). For example, F_6 can represent an N-cycle of 6 subdivisions

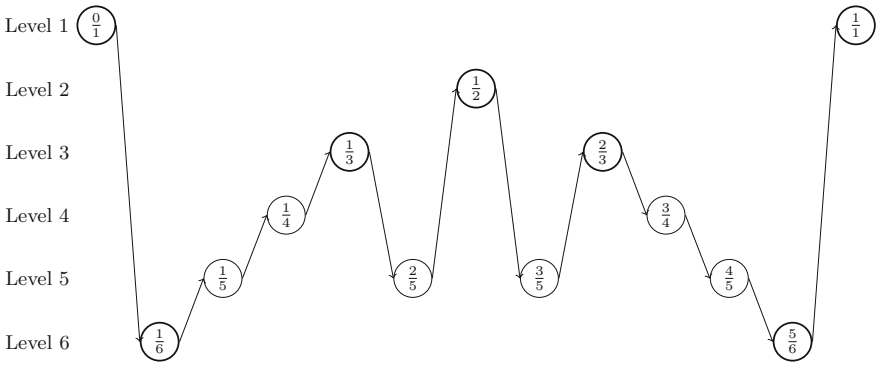


Fig. 7.9 Farey sequence F_6. Fractions a/b are placed on their respective bth level

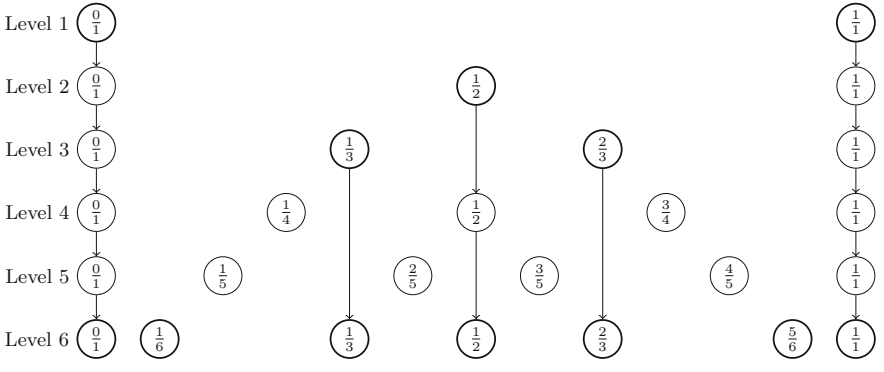


Fig. 7.10 Graphic derivation of meters 3/4 and 6/8 from Farey sequence F_6. The 2nd level is the beat level for the 6/8 meter. The 3rd level is the beat level for the 3/4 meter. Both meters have the potential to form hemiolas by placing accents on the beats of the neighbouring levels 2, 3 or even 4. The 6th level is the isochronous level of pulsation, or N-cycle

in a bar. The resulting 6 IOIs of duration $\frac{1}{6}$ can either be interpreted as a meter with 2 beats and 3 subdivisions (time signature $\frac{6}{8}$), or they can be interpreted as a meter with 3 beats and 2 subdivisions (time signature $\frac{3}{4}$). This interpretation depends on how the musical context is structured and perceived. The 2 beat cycle is marked by the fractions $\{\frac{0}{1}, \frac{1}{2}\}$ and the 3 beat cycle is marked by the fractions $\{\frac{0}{1}, \frac{1}{3}, \frac{2}{3}\}$. The Farey sequence is therefore able to represent the hierarchic structure of isochronous meters (London's I-meter). These are metrical structures that have isochronous beat periods. Meters with varying beat periods are called non-isochronous meter (NI-meter). Note that only divisiors of n can form a beat-cycle and only this relationship between N-cycle and beat-cycle generates an I-meter. Therefore, prime-numbered N-cycles will always result in an NI-meter but can never become an I-meter. In order to model an NI-meter on the basis of F_n, with a metrical subdivision by n as the N-cycle, one would use integer partitions of n to mark the starting points of beats within the N-cycle. The implementation of the above metrical hierarchies uses a

subdivision filter in order to extract all members of F_n that are related to the same level of subdivision, for example the subdivision by 6, shown in Fig. 7.10, extracts from all F_n, with $n \geq 6$, the fractions $\{\frac{0}{1}, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, \frac{1}{1}\}$. The common IOI between these onsets is of course $\frac{1}{6}$.

7.4 Conclusion

The aim of this chapter was to show that musical rhythm and meter can be modelled by using filtered Farey sequences. The Farey sequence is particularly useful because it can be scaled into different ranges of human temporal perception and varying temporal dimensions of musical form. A filtered Farey sequence can model beat subdivisions and patterns of musical meter. It is possible to analyze larger musical structures, as for example Igor Stravinsky's *Procession of the Sage*, see Sect. 7.3. Or, we can chain Farey Sequences together to form a single filtered Farey sequence of higher order, see Sect. 7.2.

The Farey tree, which is a branch of the Stern–Brocot tree, is also used for the analysis of mode-locking behaviour in weakly coupled oscillators. Large and Kolen (1995) studied their use for beat-tracking.

We are using the Farey sequence in a much broader sense and cover many aspects of musical timing, rhythm and meter. We have shown that we can arrive at a filtered Farey sequence by converting arbitrary sections of notated rhythms. This process is reversible. For example, any sequence of onset data can be normalized and then converted into a filtered Farey sequence.

The connection between perceptual timing thresholds and musical meter can be explained through so-called metrical tempo-grids (London 2012). We have demonstrated that it is possible to model these metrical tempo-grids with the use of an algorithm that is based on the Farey sequence F_{27} filtered by 3-smooth numbers. The algorithm can be easily extended in order to include higher prime factors for beat subdivisions and musical meters, see Sect. 7.3.7 and Table 7.8. We have also shown that a Farey sequence can graphically describe metrical hierarchy. Thus, a Farey sequence can generate the dot notation commonly used for metrical analysis, see Fig. 8.1, page 114, for example.

We are able to show that rhythms and meters, as well as large sections of a musical piece in notation can be analyzed and explained by filtered Farey sequences. We gave a wide range of examples from different cultures. Polyrhythms from the simple 3:2 proportion of the *hemiola* to much more complex examples can all be found in Farey sequences. The analysis of Greek verse rhythms, proportional canons from the Renaissance, African Drumming, Messiaen's non-retrogradable rhythms, any rhythm encoded as an integer sequence, highly ornamented Baroque music, in all of these cases we can operate with Farey sequences and we are thus able to generalize the method of filtered Farey sequences for music analysis. In addition, the analysis of a polyrhythmic passage from Stravinsky's *Rite of Spring* could also show a link between Stravinsky's own recording of the piece and perceptual timing limits for rhythm and meter.

References

Arom, S. (1991). *African polyphony and polyrhythm*. Cambridge: Cambridge University Press. (ISBN 052124160X).

Barlow, C. (1984). *Bus journey to parametron (all about* çoğliuotobüsişletmesi) (2nd ed., Vol. 21–23). Cologne: Feedback Verlag.

Berndt, B. C. (Ed.). (1994). *Ramanujan's notebooks* (Vol. 4, pp. 52–54). Berlin: Springer. (ISBN 0-387-94109-6).

Blecksmith, R., McCallum, M., & Selfridge, J. L. (1998). 3-Smooth representations of integers. *The American Mathematical Monthly*, *105*(6), 529–543. https://search-proquest-com.ezproxy.uleth.ca/docview/203734441?accountid=12063 (Copyright - Copyright Mathematical Association Of America Jun/Jul 1998; Last updated - 2017-11-08; CODEN - AMMYAE).

Boenn, G. (2007). Composing rhythms based upon Farey sequences. In I. Gibson, A. Stansbie, & N. Stavropoulos (Eds.), *Proceedings of the Digital Music Research Network Conference 2007* (p. 4). Leeds: Leeds Metropolitan University.

Calkin, N., & Wilf, H. S. (2000). Recounting the rationals. *The American Mathematical Monthly*, *107*, 360–363.

de la Motte, D. (1981). *Kontrapunkt. München: dtv/Bärenreiter*. (ISBN 342304371).

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*(3), 341–365.

de Vitry, P. (1964). In Reaney, G., Gilles, A., & Maillard, J. (Eds.), *Philippe de Vitry: Ars Nova*. Holzgerlingen: Hänssler Verlag.

Graham, R. L., Knuth, D. E., & Patashnik, O. (1994). *Concrete mathematics* (2nd ed.). Wesley: Reading, Massachusetts. (ISBN 0201558025).

Hajdu, G. (1993). Low energy and equal spacing. The multifactorial evolution of tuning systems. *Interface*, *22*, 319–333.

Hardy, G., & Wright, E. (2008). *An introduction to the theory of numbers* (6th ed.). Oxford: Oxford University Press.

Khan, S. A. (2012). Farey sequences and resistor networks. *Proceedings - Mathematical Sciences*, *122*(2), 153–162. https://doi.org/10.1007/s12044-012-0066-7. (01 May, 2012).

Large, E. W., Almonte, F. V., & Velasco, M. J. (2010). A canonical model for gradient frequency neural networks. *Physica D*, *239*(12), 905–911. https://doi.org/10.1016/j.physd.2009.11.015.

Large, E. W., & Kolen, J. F. (1995). Resonance and the Perception of Musical Meter. In Griffith, N. & Todd, P. M. (Eds.), *Musical networks: Parallel distributed perception and performance* (pp. 65–96). Cambridge, Massachusetts: MIT Press. (ISBN 0262071819).

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.

Lerdahl, F., & Jackendoff, R. (1996). *A generative theory of tonal music (Reprint of 1983 ed.)*. Cambridge, Mass.: MIT Press. (ISBN 9780262621076).

London, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Marron, E. (1990). *Die Neue Rhythmik-Lehre*. Brühl, Germany: AMA Verlag. (ISBN 979-0700136604).

Martelli, A., Ravenscroft, A. M., & Ascher, D. (Eds.). (2005). *Python cookbook* (2nd ed.). Sebastopol, California: O'Reilly Media Inc. (ISBN 0596007973).

Messiaen, O. (1995). *Traité de Rhythme, de Couleur et d'Ornithologie* (Vol. 1). Montrouge: Alphonse Leduc. (ISBN 978-004628892).

Moore, F. R. (1990). *Elements of computer music*. New Jersey: Prentice Hall. (ISBN 0132525526).

Papadelis, G., & Papanikolaou, G. (2004). The Perceptual Space Between and Within Musical Rhythm Categories. In Davidson, J. W. (Ed.), *The music practitioner : Research for the music performer, teacher, and listener*. Burlington: Ashgate. (ISBN 0754604659).

Pratihar, S., & Bhowmick, P. (2017). On the Farey sequence and its augmentation for applications to image analysis. *International Journal of Applied Mathematics and Computer*

Science, *27*(3), 637–658. https://www.degruyter.com/view/j/amcs.2017.27.issue-3/amcs-2017-0045/amcs-2017-0045.xml?format=INT.

Schroeder, M. (1991). *Fractals, Chaos, power laws: Minutes from an infinite paradise*. New York: W. H. Freeman and Company. (ISBN 0-7167-2136-8).

Sloane, N.J.A. (2011). *The Stern-Brocot Tree*[Web page]. World Wide Web electronic publication. http://oeis.org/stern_brocot.html (Retrieved March 2018).

Stravinsky, I. (1960). *The rite of Spring*. Columbia Records. https://www.discogs.com/Igor-Stravinsky-Columbia-Symphony-OrchestraNew-York-Philharmonic-Orchestra-Stravinsky-Conducts-Le-Sa/release/10969649 (Columbia Symphony Orchestra).

Stravinsky, I. (1967). *The rite of Spring* Vol. (HPS 638). London: Boosey & Hawkes. (Re-engraved edition, plate B.&H. 19441. First performance 1913. Revised 1947.).

Taylor, S. A. (2012). Hemiola, maximal evenness, and metric ambiguity in late Ligeti. *Contemporary Music Review*, *31*(2–3), 203–220. https://doi.org/10.1080/07494467.2012.717362.

Tomás, R. (2014). From Farey sequences to resonance diagrams. *Physical Review Accelerators and Beams*, *17*, 014001. https://doi.org/10.1103/PhysRevSTAB.17.014001.

Wannamaker, R. (2012). Rhythmicon relationships, Farey sequences, and James Tenney's spectral canon for conlon nancarrow (1974). *Music Theory Spectrum*, *34*(2), 48–70. https://doi.org/10.1525/mts.2012.34.2.48.

Chapter 8
Models of Musical Meter, Temporal Perception and Onset Quantization

8.1 Introduction

In this and the following chapter I will describe how it is possible to extract rhythm data from audio recordings, and to transcribe note onsets into music notation. For the collection of note onset data used by the rhythm quantization and transcription algorithms, which I develop in Chap. 9, we need to understand and apply current techniques of onset detection from various data sources. For instance, we would like to know how accurate they are, and how much noise is likely to occur in the data when those methods are being used.

Important aspects of temporal perception of musical phenomena are taken into account for the design of our quantization model. We will also analyze the relation between a notated score and its musical performance, which always involves expressive timing, sometimes also called *agogics*. Musical ornaments give performers many opportunities to use expressive timing. Therefore, I have chosen a richly ornamented piece from the repertoire of Baroque keyboard music, the *Aria* of the *Goldberg Variations*, in order study various challenges for music quantization and transcription. Because this music is based on the Western notion of musical meter, I will start with a description of meter and analyze its various implications on the perception of rhythm.

Gestalt theory offers viable approaches for the analysis of musically performed durations. One key element of our model uses a special method of grouping and classification of durations. I have used clustering methods in the past (Boenn 2011), but now I can present a new method that is both more efficient and easier to use.

Various results from research into musical meter, rhythm analysis and perception are presented. Finally, we will look at the current state-of-the-art of quantization and tempo-tracking.

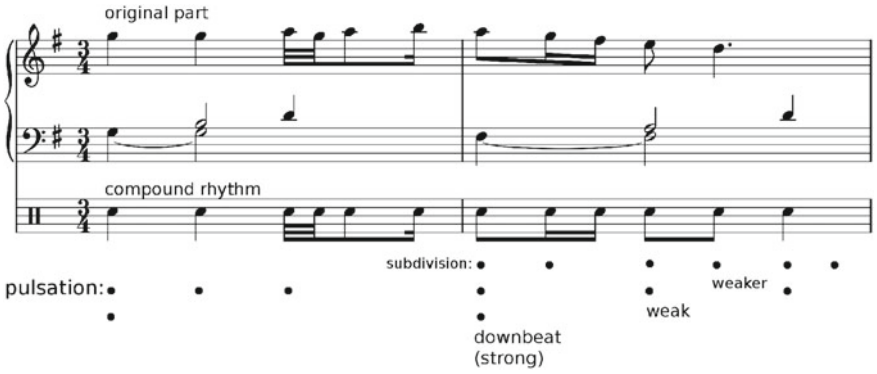


Fig. 8.1 Beginning of the *Aria* of J.S. Bach's *Goldberg Variations*. Rows of equidistant dots denote pulsations on a specific level within the metrical hierarchy introduced by the nature of the 3/4 meter. The upper row represents 1/8 note pulsation, the centre row represents the beat level in 1/4 notes, and the lowest row indicates the strong downbeats, which form a pulsation in dotted 1/2 notes. Notes are found on weaker metrical levels if they do not coincide exactly with the downbeat

8.2 Musical Meter

We have discussed musical meter in the context of filtered Farey Sequences in Chap. 7, Sect. 7.3.7, and I introduced a measure for metric complexity of musical phrases in Chap. 4, Sect. 4.5. In this chapter, I will revisit the concept of meter, and we will learn more about its practical application, as well as about the perceptual aspects of meter and timing.

One of the main conventions in notation is to align note durations with a hierarchic metrical grid, i.e., a grid composed of durations. Pulsation and meter are sometimes referred to as the simplest form of rhythm (Arom 1991). Meters are indicated by time signatures, which are also integer ratios relating to the whole note as a reference. The purpose of a time signature is to define the duration of a bar and how it is composed out of smaller units of durations. The basic metrical grid of beats forms within a bar and it is usually continued for the whole length of a section of music. The first beat in a bar is called the *downbeat*. The following Fig. 8.1 illustrates how note durations of a musical piece relate to a metrical grid that is defined by the time signature *and* by the compound rhythmic surface of the music. Heavy downbeats, weaker beats and further subdivisions of the beat are represented by particular rows of dots below the piano staves. The interplay between strong and weaker metrical positions in time is a feature of metrical hierarchy. The musical structure can contradict this regular scheme by placing accents on weaker parts of the meter, a process known as syncopation. A single dot denotes onsets of a particular metrical level – a representation that is widely used in music theory (Lerdahl and Jackendoff 1996; London 2012). Although a beat carries a duration in the score, there is no direct indication of the actual tempo nor of absolute physical time in notation. This might have changed with the advent of the metronome in the 19th century, but there are still problems. A metronome is a

mechanic oscillation, which contradicts the usual ebb and flow of expressive timing in music. Although composers could include absolute timing information in their scores, the metronome value of the beat can only give an *indication* of the tempo in a unit called beats per minute (BPM).

Notation uses fairly simple ratios in order to represent musical durations and rhythmic structures. It was also found in experiments that humans categorize groups of durations according to their vicinity to simple ratios and maps of these categories can be found for example in Desain and Honing (2003). Western music theory often discusses rhythmic composition techniques on the basis of the ratios that are being notated within Western notation, for example Stockhausen (1988). In this article, Karlheinz Stockhausen draws parallels between rhythmic textures, tempo relations and pitch intervals of the overtone series in order to have a unified serial approach to pitch and rhythm composition.

8.2.1 Necklace Notation of Rhythm and Meter

London (2012) and Sethares (2007) take advantage of necklace notation for musical rhythms and metric hierarchies. This is an early form of notation, which dates back to 13th century Arab music theory, when Safī al-Dīn al-Urmawī published his *Kitāb al-adwār* ('Book of Cycles') (al-Dīn al Urmawī 1980; Sethares 2007). The tenet of modern theory is that musical meter is a cyclical repetition of patterns of attentional energy focused by the listener on specific moments within a musical meter. A cyclical representation supports the notion of repetitive patterns. Although music often works with the principle of variation, any variations of rhythmic patterns are carried out on a backdrop of culturally accepted and individually learned metrical patterns. London (2012) calls this process *entrainment* of metrical patterns by the listener. Small changes of musical parameters will keep the listeners attention while the inherent stability of the underlying metrical hierarchy helps to fully appreciate the changes and to grasp their aesthetic meaning. The following figures depict such stable metric patterns from different musical cultures. A circle represents a note onset event as well as an expectancy peak in an entrained pattern of musical meter. In necklace notation, time flows continuously and clockwise around the circle. Figure 8.2 shows the Arab rhythm *thāqil thāni*. Large dots indicate the initial time or foot that is always sounded, medium dots denote medial time units, optionally sounded, whereas the smallest dots indicate the final time unit that is always silent (Right 2001). Figure 8.3 shows the Ghanaian *Ewe* rhythm. Variations of this rhythm appear in other African cultures as well, for example in Central Africa and in Nigeria. The difference between them is that their particular version starts at a different time point within an otherwise identical rhythmic pattern.

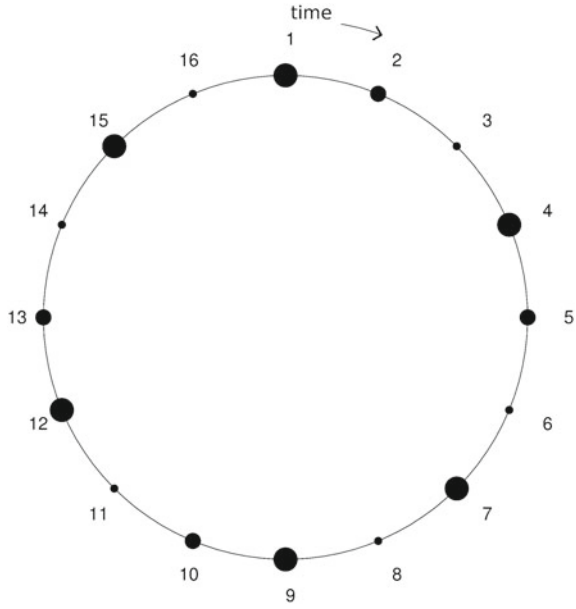


Fig. 8.2 The Arab rhythm *thaqil thani*

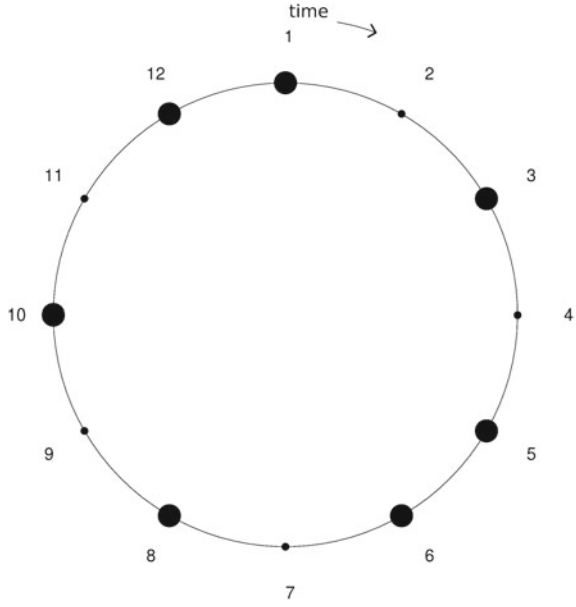


Fig. 8.3 The *Ewe* rhythm from Ghana, Africa

8.2.2 *Meter and Entrainment*

Meter can be described as an attentional dynamic process, whereby listeners actively anticipate future events using their individually entrained metrical framework, which is resonating to musical rhythms (Large and Jones 1999; Large 2001; London 2012). Large (2001) describes meter as a self-organising dynamic structure based on a network of self-sustaining oscillations. Here, meter perception emerges from the formation of neural patterns stimulated by external musical rhythms. The oscillator-based entrainment model of Large and Palmer (2002) predicts the categorisation of temporally changing event intervals into discrete metrical categories. London (2012) regards meter as a form of entrainment, which enables listeners to extract temporal invariants from the music. The main structure of those invariants is encapsulated in the framework of metrical hierarchies and patterns of temporal expectancies that a human being learns when exposed to music over the course of his or her life.

Meter and the Dynamics of Attending

In many Western and non-Western musical practices, meter serves as the foundation for the performance and perception of rhythmic figures. But, meter is not just a neutral background structure, and its purpose is not only to enable the synchronization of rhythmic gestures. Both meter and rhythm perception mutually influence each other. "Just as rhythmic patterns evoke metric responses to the listener, at the same time the listener's metric entrainment can give rise to a rhythmic figure" (London 2012, p. 75).

Meter is a "stable, recurring pattern of attentional energy". This pattern is represented as a cycle. Often meter is a metacycle, that is a set of hierarchically coordinated component cycles (London 2012, p. 80). Meter is a dynamic system which has various modes of excitation or resonance. The modes of resonances stem from different metrical levels and they form together a hierarchy of expectation (Large and Jones 1999).

A musical meter very often contains more than one layer of pulsation (Barlow 1991). Entrainment of metrical hierarchies can be modelled by Gaussians centered around metrical time points on different layers. The superposition of multiple layers will automatically create higher peaks for strong beats, the highest peak for the downbeat, and lower peaks will represent the weaker beats. The appearance of beats too early in a performance can then still be mapped to the expected beat position because of the continuous valued curve representing the level of expectancy that a listener is attuned to Large and Palmer (2002). While a weak beat has lost its influence on the current expectancy level the downbeat has already taken over and dominates the listener's expectancy within a certain window of time until a weak beat's expectancy rises again (Large and Palmer 2002, p. 12, Fig. 3).

In general, meter comprises a measure level starting with the downbeat, a tactus or beat level, which divides a bar into distinct timing regions, and often there are one or more levels of subdivisions per beat. There is no need for a distinction between meters and so-called hyper-meters, which may span across many consecutive bars,

because the number of metrical levels usually varies significantly during the course of a musical composition. The periodic change of attentional energy in meter perception leads to the assumption that metrical accents are generated by the listener. They are based on the strength and focus of entrained temporal expectancies (Barlow 1991; Huron 2006; Large and Jones 1999; London 2012). Music performers and composers can play with these expectancies, either to reinforce them, or to override them more or less gently.

The physical durations of meters, beats and subdivisions are chosen for musical reasons as well, and in real performances they always arise out of the musical and acoustic context. The perception of rhythm and meter, and thus their performance, is guided by certain periodicities and temporal limits. The temporal range for metric entrainment is between 100 ms and 5 or 6 s. A sense of beat is given in the range between 200 and 250 ms to 2 s. There is an attentional peak or preference for time spans around 600–700 ms, the so-called indifference interval (we will come back to these limits in Sect. 8.3). Those perceptual constraints lead to a specific grid of metrical and rhythmic possibilities (London 2012, p. 45, Fig. 2.6.) see also Table 7.8 on p. 108.

One can see that a chosen tempo for the central beat value automatically excludes certain note durations and lengths of meters because they exceed corresponding perceptual timing limits. Different tempi lead to specific sub-trees of metrical possibilities. Tempo changes on the tactus level affect the perception of meter and these circumstances lead also to perception of groups and meta-groups in the performed rhythms.

There are certain rhythmic patterns that can theoretically fit in more than one metrical hierarchy. London (2012) calls them "metrically malleable patterns". But set into the context of a musical situation, there can only be one way of perceiving malleable patterns. Any pattern will be performed in a specific manner that unambiguously belongs to a specific meter due to the articulations employed by the performer. Musical articulation consists of phrasing, grouping, placing accents, and so forth, and it is metrically informed (Large and Palmer 2002).

The perception of polyrhythms can also be influenced by the articulation of the performer. They are created because of specific integer ratios between note durations. Let us take for example the *hemiola*, two voices with durations in the ratio of 2:3, see Fig. 8.4. The resulting compound rhythmic structure can be played with at least two different ways of articulation. The first one accentuates the onset at $1/2$, the second will have the onsets $1/3$ and $2/3$ standing more in the foreground but still with the main accent on $0/1$ and $1/1$. One can produce polyrhythms easily within a metrical structure because of the basic isochronous pulsation within a meter, see also Sect. 7.3.2, Table 7.2, p. 93. The system of notation supports polyrhythms, because it is based on the concepts of metrical hierarchy. The interaction between metric patterns and rhythmic figures creates many possibilities for performers and composers to play with listeners' expectations. Performers, for example, are able to phrase and place accents in contrast to the underlying metrical framework in order to communicate a specific musical meaning and to evoke specific listener responses.

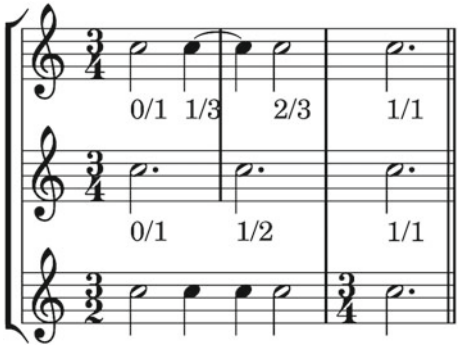



Fig. 8.4 Hemiola with onsets marked by small integer ratios. The lowest staff shows its compound rhythmic structure

London (2012) developed a circular representation of meter, similar to the necklace notation we have seen earlier, but extended through the use of subcycles, which reveal the different levels of a metrical hierarchy. This kind of representation underlines the cyclical nature of musical meter, which is grounded in the recurrence of patterns of attentional energy. At the lowest level of a meter is the N-cycle, a set of N attentional pulses with the shortest period in a meter, but not shorter than 0.1 s. N is always an integer so that there can be 8-cycles, 9-cycles and so forth. All higher levels of a meter, i.e., the subcycles, can be constructed by taking integer multiples of the period of the N-cycle, but there are constraints for this procedure. The subcycle carrying the tactus or beat of the meter is called the beat-cycle. Listeners have a perceptual preference for beat periods between 0.5 and 0.7 s. For N-cycles with N $>= 4$ there can be a special cycle marking approximately one half of a measure with an additional attentional period. The N-cycle and all subcycles viewed together form the metrical type. There are certain ranges for the tempo of a metrical type, where the periods of some individual subcycles align with the psychological timing limits mentioned before. These tempo ranges define so-called tempo-metrical types.

Finally, metrical hierarchies and patterns of expectations are not solidified but temporally elastic structures, which exhibit variance at different time scales due to the use of agogics in musical performances. The *Many Meters Hypothesis* states that "a listener's metric competence resides in her or his knowledge of a very large number of context-specific metrical timing patterns. The number and degree of individuation among these patterns increases with age, training, and degree of musical enculturation" (London 2012, p. 182).

8.3 Temporal Perception

The following perceptual timing thresholds have important implications for the analysis of musical rhythm and meter. They will necessarily influence our computational model of musical rhythm in terms of the categorization of rhythmic timing informa-

tion, for example in algorithms that analyze sets of inter-onset times. However, one should treat the thresholds as guidelines not as absolute values. The physical timing values reported from various experiments have to be taken *cum grano salis*. There are inter-subjective differences, averaging procedures across subject groups and a mixture of experimental stimuli have been used, sometimes based on real musical performances using different styles, but often laboratory generated stimuli have been used. Nevertheless, these thresholds provide a highly valuable knowledge base for the analysis of real musical performances. They are helpful in providing a psychological foundation for the analysis of musical meters and their relationship with the performance and perception of musical rhythms.

Experiments have shown that human listeners spontaneously group identical and isochronous stimuli into groups of twos or threes. In other words, listeners are having a sense of accent on every second or third stimulus although such accents are not present in the stimulus. This effect of binary or ternary grouping is known as *subjective rhythmization* (Bolton 1894; Parncutt 1994).

In the past, researchers have proposed certain ranges of perceptually relevant beat periods (London 2012; Parncutt 1994; Warren 1993; Westergaard 1975), which are remarkably similar. Following the thorough review of the research on tempo and beat perception by London (2012), musical beat periods are found in the range from 0.2 s up to 2 s, or 300 to 30 BPM. If one takes into account binary subdivisions at the fastest end, as well as ternary meters divided by the slowest beat, then one finds an outer timing envelope for metrical entrainment that ranges from 0.1 s to somewhere in-between 5 and 6 s. Within this range there is another significant perceptual threshold relevant for the analysis and performance of musical rhythms. Parncutt (1994) had subjects tapping to a variety of rhythmic patterns. He found a peak of maximal pulse salience at 0.6 s within a general range of pulse perception between 0.2 and 1.8 s. Durations between 0.6 and 0.7 s are also known as *indifference intervals* (Fraisse 1963; Wundt 1911), which are neither overestimated (judged too long) nor underestimated (judged too short) by subjects in various studies. At 600 ms, on average, lies also the spontaneous tapping period (Fraisse 1999) although this value seems to vary with the age of the subjects. London (2009) calls the period of 0.6 s *tempo guisto*, by sheer coincidence it corresponds to 100 BPM. London (2009) also suggests that this period relates more strongly to the neurobiological nature of motor control and kinematic movement as in walking and running, clapping hands (Repp 1987), and so forth, as opposed to metabolic periodicities, such as breathing and heart beat.

8.3.1 Shortest Timing Intervals

Hirsh (1959) found that a 2 ms separation is required in order to discern two tone onsets, otherwise they would be heard as a single one. But, if the listener is asked to report the order of the two onsets correctly, then they need to have at least a 20 ms interval between them.

It appears that fast played ornaments and gestures, such as trills, drum rolls, arpeggios or bow tremoli, fuse perceptually together into a single sound event. The time periods for this kind of musical gesture are approximately 70–90 ms long. The upper limit of haptic frequencies is 11 Hz on average, which corresponds to a period of approximately 91 ms (Verplank et al. 2000). Trills of professional piano players have been analyzed by Brown and Smaragdis (2004). The highest frequencies can reach up to the range of 14 Hz (peak values could even reach 16 Hz), with a period of about 71 ms, i.e., the time elapsed between individual notes of the trill. It is interesting to note that such periods of oscillation stay just above the human limit of pitch perception: 16 Hz ≡ 62.5 ms.

8.3.2 The 100 ms Threshold

London (2012) found that only above the 100 ms threshold elements of rhythm and meter can be recognized one at a time: "The lower limit for meter, that is, the shortest interval that we can hear or perform as an element of rhythmic figure, is about 100 ms". This threshold is consistent with music psychological research. For example, the lower limit for subjective rhythmization is 115 ms (Bolton 1894). The threshold for reliable durational discrimination lies also at 100 ms (Hirsh et al. 1990). It is the minimum period, which is necessary for the cortical processing of sound patterns (Roederer 2008). It is also the fastest period of pulsation to which subjects could synchronize a four times slower tapping period (Repp et al. 2002). However, the average of his experiments peaks at 120 ms. 100 ms is the shortest duration in ride cymbal patterns of jazz drummers (Friberg and Sundström 2002). The fact that 100 ms is the shortest possible metric subdivision (London 2012) has consequences for the durations of performed metrical structures as we will see in Chap. 9.

8.3.3 Fastest Beats

Beats are the governing force within a metrical hierarchy. Given the minimum duration of 100 ms for any rhythmic element that has an identity of its own as part of a musical gesture and given the assumption that "hearing a beat requires at least the potential of hearing a subdivision", (London 2012), the threshold for the fastest beat must be assumed to be in the range of 200–250 ms. This range correlates nicely with various results from music psychological experiments. It is the borderline period between holistic versus analytic processing of rhythmic events (Michon 1964). It is supposed to be the limit for short-term memory (Crowder 1993). It is the cutoff for backward masking (Massaro 1970). In this range a shift happens in nature of the just noticeable difference (JND) for duration (Friberg and Sundberg 1995). For (Westergaard 1975), this is the range of the fastest beats: 240–300 BPM. Repp (2002b) found this rate as the limit of the subdivision benefit. There is a benefit for listeners to tap along at certain subharmonic periodicities, e.g., 1:2, 1:3 or 1:4, if the stimulus pulsation is sufficiently fast. Warren (1993) found here the upper limit

for melody recognition. It is also the upper end of range of pulse salience (Parncutt 1994).

8.3.4 Slowest Beats

On the other side of the beat spectrum we find the slowest tempo range of 40–30 BPM, or durations of 1500–2000 ms per beat (London 2012; Westergaard 1975; Warren 1993). Here one finds a correlation with the upper limit of subjective rhythmisation (Bolton 1894; Fraisse 1999) and a shift from anticipatory attending to reaction time (Woodrow 1932). The maximum time period for a ternary meter would then correspond to 4.5–6 s, which is equal to the limit of psychological present (James 1890, 1950; Pöppel 1972; Michon 1978; Fraisse 1984).

8.3.5 The Perceptual Time Scale

Based on the perceptual timing thresholds presented, we can show them all together mapped to the basic musical time structures of beats, subdivisions and bars, see Fig. 8.5. Interestingly, the average indifference interval of 650 ms sits at the centre of the logarithmic scale of musical beat periods.

8.4 Onset Detection

The detection of note onset times in musical performances is crucial for the creation of experimental data sets. Previous research about tempo tracking and automated transcription used onset data from various data sources, like manual tapping, audio and MIDI recordings. They also used different musical styles. We expect that the detection and extraction of onset data will introduce noise. The question is whether the noise can be ignored for the purpose of rhythm quantization. Therefore, we need to know more about the kind of errors and their approximate levels that are going to be introduced by the various methods of measurement.

8.4.1 Manual Tapping

When using a computer keyboard for manually tapping along an audio signal, Wright et al. (2004) reported latency and sometimes significant jitter from the keyboard. The best results of low latencies and jitter were achieved when the ALSA and JACK audio drivers had been set at low latencies on a Linux system configured for real-time audio.

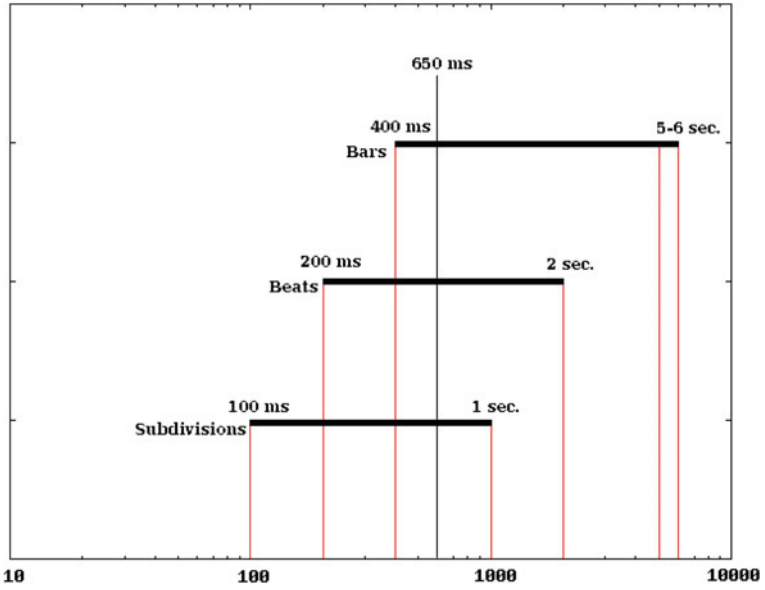


Fig. 8.5 Perceptual timing thresholds and musical time structures

MaxMSP under MacOS X achieves comparable results when the "overdrive" option is turned off. Under those optimized conditions one has to take on average 21–22 ms of latency from computer keyboards into account. A quantization algorithm has to be robust enough in the presence of jitter and should run preferably under a system configured properly for real-time audio. 22 ms latency is equivalent to the point-of-view of a musician controlling a sound source that is 7.55 meter away. The maximum rate for repetitive musical timing gestures, e.g., trills, is on average 14 Hz, which means that their components are approx. 71 ms long (Brown and Smaragdis 2004). The fastest periods on average for purely rhythmic events is around 100 ms. Due to receiving acoustic feedback of their action we expect some influence of the latency on the player's performance at very fast tempi. Professional musicians using the manual tapping apparatus could perhaps adapt in a similar way to their adaptation to large room acoustics.

A different approach that yields a much better latency is the use of a force-sensitive resistor (FSRs) built into a voltage-divider circuit, and connected to a micro-controller, such as the Arduino, for example, and from here connected to the computer via USB port. Our tests have shown that the latency for sending a 4-byte integer from an Arduino Mega at 9600 baud to a laptop is approximately 3.5 ms (plus or minus 0.5 ms) on average.

8.4.2 Onset Data Extracted from Audio Signals

Over the last two decades, certain standards of audio feature extraction have emerged, like MPEG-7 for example Martínez (2005). An overview of signal processing methods for audio feature extraction is given in Peeters (2004) as part of the European CUIDADO project coordinated by IRCAM. Also, the *Mazurka Project* developed their own methods and tools, notably VAMP plug-ins written by Craig Stuart Sapp for the Sonic Visualiser. We also mention the Libextract library by Jamie Bullock that provides around 50 spectral and other features from audio data. The Aubio library by Brossier (2006) has mainly been used for our development in conjunction with a sample editor for hand-editing of the markers. Onset detection still poses some problems notably when it comes to analyzing vocal music and also string instruments. Sometimes small clusters of markers occur around a single audio onset or some of the audio onsets are not detected and markers are missing. Usually, one adjusts the threshold parameter settings on the basis of test runs of the onset detection, but hand-editing of the resulting markers is still necessary if one needs to obtain a correct outcome of the tracking process. This is the case in our study with its main focus on quantization and transcription of the collected onset markers. An in-depth survey of onset detection algorithms can be found in Collins (2006).

Recently, Grosche and Müller (2011) published a new method to extract mid-level local pulse information from audio recordings. Their method uses a compressed magnitude spectrum, on which basis they perform a calculation of a temporal derivative to arrive at a so-called *novelty curve*. A spectrum analysis of the novelty curve generates maxima for the selection of appropriate sinusoidal kernels, which accumulate in an overlap-add synthesis. The half-wave rectified output is called the *Predominant Local Pulse*. Test results show that this is a promising approach for capturing local tempo information in Classical music. A comprehensive account of recent spectral audio analysis techniques with many applications to music analysis has been published by Müller (2015).

8.4.3 Adjacent Interval Spectrum

The analysis of performed musical durations is one of the key tasks of our proposed method of quantization, which allows to capture subtle tempo changes and expressive timing in human performances. To achieve this, it is very helpful to visualize musical performance data by using an adjacent interval spectrum, an original method developed by Kjell Gustafson (Toussaint 2004). The main idea is to plot timed durations as boxes against the vertical axis, and the sequential order of their occurrence against the horizontal axis, like for example in Fig. 8.6. With an interval spectrum it easy to compare the plotted durations with each other. In a next step, one has to detect quasi equal durations, which, despite local variations, belong to the same kind of duration in the original score. In Fig. 8.6 for example, one can see that the

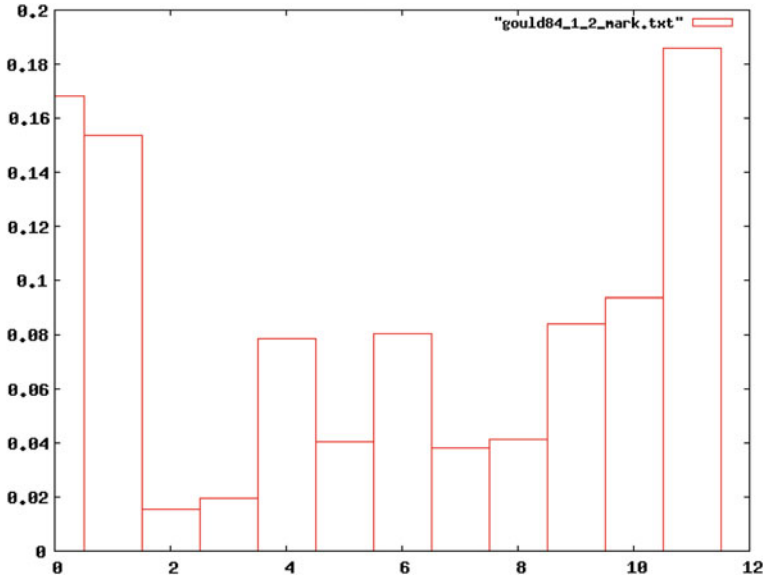


Fig. 8.6 Adjacent interval spectrum of the compound rhythm of the first two bars of J.S. Bach's *Aria* performed by Glenn Gould in 1981, with boxes plotting the normalized durations of inter-onset intervals on the y-axis. Figure 8.1 shows the score

durations number 4, 6, 9 and 10 have values close to 0.08. There is also a significant distance between them and other durations. But, over the course of these two bars they seem to become slightly longer each time. A comparison with the score in Fig. 8.1 can then reveal how a particular musical event has been timed by Gould and how it was put into a musical context, for example a tendency to slow down with a phrase ending at the end of bar 2 is reflected across various duration classes. One can also observe in Fig. 8.6 that the deceleration had already started at the beginning of bar 2 with duration number 6. The last duration 11 seems to stand out from the others but it is according to the score the same quarter note value as the durations 1 and 2 from the beginning of the piece. This deliberate bending of musical time is a well-known means of human expression, and a wide-spread practice of performers of Western classical music, but not only here. The Indonesian Gamelan and the Japanese Gagaku also use continuous tempo changes. And, although musical tempo seems to fluctuate all the time, a human listener seems to have no problems to follow these tempo changes, for example to tap the foot along, and one can take a delight in the bending of time as an important part of the musical experience. But, for the design of an automated quantizer and transcription system, it is still a challenging task to be accurate when trying to follow tempo changes.

8.4.4 Is Knowledge of Onset Times Sufficient?

One might ask why we can be so sure that onset data alone is sufficient for successful rhythm transcription. Is there enough information about the perceptual qualities of rhythm and meter encoded by a stream of inter-onset intervals? It might be useful to look at the psychoacoustics of sound perception. The sound of musical instruments shows transients that are critical for the ability of humans to discern different musical instruments. Such transients are usually very short and happen during the attack time and so the note onset coincides with a very important physical and psychoacoustic feature of musical sound. It seems that the very first part of the sound, which occurs only briefly as the transient, literally earmarks the moment in time when a new sound starts. One might think of a way how compound rhythms are perceived by a listener, namely through the connection of the individual attack points of different voices in the music.

Other researchers discovered a great sensibility of the nerve cells of the ear to detect onset and offset times of musical instruments playing a sequence of musical events, as pointed by Brossier (2006):

> Another type of temporal encoding operated by the human ear allows for the analysis of sonic events: some of the nerve cells are triggered at onset and offset times, where a sound starts and finishes (Whitfield 1983). The attack of the note, where the sound starts rising, triggers these nerve cells, while the sustained part of the sound, where the note is held, does not. A detailed analysis of this phenomenon was given in a study of the perceptual attack time (Gordon 1984). The study included perceptual tests in which listeners were asked, while listening to two different sounds separated by different time delays, to press a button if the attacks of both sounds were perceived as simultaneous. Gordon could measure accurately the perceptual attack times and found that these times were consistently different amongst different listeners. The tests showed that perceptual attack times of tones could be perceived significantly later than the physical onset of the sound in the music signal, up to a few tens of milliseconds, depending on the instrument played and the way it is played. Gordon (1984) observed that the perceptual attack time of several sounds was dependent on the timbre of the instrument, this quality of a sound which enables us to distinguish one instrument from another (Grey 1975).

Human perception of transients during attack time are crucial for the source identification of musical sounds (Iverson and Krumhansl 1993). Experiments where transients have been cut off have shown that instrumental timbre becomes hard to identify, this applies in particular to the family of wind instruments. This acute awareness of the onsets in perception also feeds back into instrumental practice. The precision of playing note onsets is usually higher than that of playing note offsets (Kilian 2004).

Playing an instrument and listening to music depends heavily on certain temporal structures in our perception. There are musical phenomena of phrasing, expressive timing, a so-called dynamic function (Thakar 1990), which describes the ebb and flow of tension in music as a whole, and there are phenomenal differences between musical performances of the same piece even when they happen under similar conditions. Central to all musical phenomena is the question of the temporal occurrence of sounds in relation to the sum of the effects of all musical events, which happened before, and in expectation of what will happen immediately afterwards as a direct

consequence. This is the causal continuity in music. A better understanding of the structure of human time perception is an important prerequisite in order to further study this question.

8.5 Agogics

The Oxford Companion to Music defines agogic accent as follows, after (Efrati 1979, p. 101):

> This term [...] is applied to that kind of accent which belongs to the nature of the phrase, as distinct from the regular pulsation of the so-many-beats in a measure, and which is produced rather by dwelling on a note than by giving it additional force. The first note of a phrase often suggests the desirability of a slight lingering, which constitutes an agogic accent. So does a note longer than, or higher or lower than, those that have preceded it. So does a pungent discord about to proceed to its resolution.

This notion of agogic accents underpins how Thakar's *dynamic function*,[1] i.e. the gathering and dissipation of musical energy, permeates many different parameters. In our opinion, it is according to the *dynamic function* that musical agogics work. Agogics are often related to rubato playing[2] for which there are at least two main historic traditions. The old tempo rubato meant to maintain the tempo of the musical accompaniment whereas the melodic voice, by contrast, was allowed to halt, to slow down or to speed up, thereby creating *notes inégales*. The resulting new relations of consonance and dissonance between the melody and the accompaniment were set intentionally and formed a part of an improvisational element where the musical texture was enriched by new appearances of dissonances and consonances. Of course those new relations needed to follow the common rules of counterpoint but they enriched the musical experience of a piece and added musical events not explicitly written down in the score, similar to the practice of adding ornamentation to a melody.

Historic theoretical works discussing the latter type of rubato are, for example, by Pier Francesco Tosi, Johann Joachim Quantz and Carl Philip Emanuel Bach.

> Slight tempo modification in the principle voice, for the sake of expression, was permitted. The so-called "tempo rubato" took its origin from seventeenth-century Italian vocal music, and was first described by Tosi in his "Opinioni" (1723). His definition: "rubamento di Tempo, sul movimento uguale d'un Basso." Whilst the principal voice was allowed some tempo variation, the accompaniment has to carry on its regular pace.
>
> C.Ph.E. Bach (Chapter XXIX, §18): "When in a polyphonic composition the bass, as well as the ripieno parts, have to hold a long note, whilst the principal voice maintains its special movement, *even introduces some irregularities now and then;* the accompanist would do well, in order to maintain a strict time, to go on striking the beats with his right hand." (Efrati 1979, p. 217)

[1] Not to be confused with terminology of musical dynamics such as *forte* and *piano*.

[2] Tempo rubato means literally 'stolen time'.

The other form of tempo rubato refers to the notion of the complete set of voices changing their tempo according to the Gestalt rule of common fate (Deutsch 1999), thus not changing the synchronization between different voices but changing the tempo as a means of expression to serve the *dynamic function* of the musical texture.

There is an understanding of rubato that demands that the tempo changes only for a short period of time and then it returns back to a sort of average value applied to the entire piece. A rule attributed to Chopin[3] states that if within one bar one slows down for example, the next bar has to compensate by accelerating proportionally to the previous bar's rallentando. In this way performers try to level the amounts of rallentando and accelerando during a piece so that these tempo deviations always return like a flexible matter into its state of rest, i.e., to the overall average tempo of the performance. Rubato playing does not have total liberty but it is set into a context: "The detailed temporal properties of performed music […are] often referred to as temporal micro-structure […]" (Clarke 1999).

The presence of rubato and expressive timing in musical performances means that automated tools for quantization and transcription need to become aware of continuous tempo changes. However, tempo itself is a perceptual construct and not easy to extract from the actual onset data. Therefore we recognize that a Gestalt approach is needed in order to cope with rubato playing and expressive timing. We will explain this concept in the next section.

8.6 Gestalt Theory

A part of our research was inspired by Gestalt-Psychology. The method of grouping inter-onset times into distinct duration classes, which I will develop in the Chap. 9, depends on the notions of proximity, similarity, good continuation and on the principle of common fate. Durations that look identical in the score are not exactly identical in a musical performance. In order to classify durations according to similarities, it was useful to see the spatial information in graphical representations of the lengths of actual note durations, see the use of the adjacent interval spectrum in Sect. 8.4.3 and Fig. 8.6.

In addition, time seems to be often described in spatial terms (Köhler 1947). Proximity is a spatial concept after all, or to quote from Köhler's "Gestaltpsychology", p. 88:

> Experienced time has certain characteristics in common with experienced space, particularly with the spacial [sic] dimension which is indicated by the words "in front" and "behind". Words which refer to relations in this dimension are used as terms for temporal relations everywhere and in all languages. In English we may have something "before" or "behind" us both in the spacial [sic] and temporal meaning: we look "forward" in space as in time, and death approaches in time just as somebody approaches us in space. From the point of view of isomorphism, one would expect that there is a corresponding kinship between

[3]Personal communication of the author with the German pianist, organist and pedagogue Elisabeth Wangelin–Buschmann.

> the physiological correlate of the temporal and that of this spatial dimension. At any rate, temporal "dots" form temporal groups just as simultaneously given dots tend to form groups in space. This holds for hearing and touch no less than it does for vision.

This passage can also be found in Efrati (1979), p. 75, where he reports views on the musical line in the context of Bach's Solo Violin Sonatas and Partitas and the Suites for Solo Cello.

In Chap. 9, I will demonstrate that the "temporal dots" are equal to the recorded inter-onset intervals. They form "temporal groups", which can be analyzed by the grouping algorithm in Sect. 9.2. It uses the concept of proximity, i.e., similar magnitudes of duration cause the formation of a temporal group. We see a specific temporal group as a duration class and it is the ratios between different duration classes that form rhythmic motives, phrases, sections, and so on. The important point is that the duration classes are not fixed arrangements of magnitudes but rather that their mean values can vary significantly in the course of a performance. Although this happens during a musical performance the ratios between duration classes remain generally stable, a fact that is linked to the other Gestalt principles, such as good continuation and common fate. Music psychologist Diana Deutsch (1999, p. 300): "One is proximity: closer elements are grouped together in preference to those that are spaced further apart." The second principle is similarity, which means similarity of form, or a similarity of one attribute of from. "A third, good continuation, states that elements that follow each other in a given direction are perceptually linked together. …A fourth, common fate, states that elements that change in the same way are perceptually linked together. As a fifth principle, we tend to form groupings so as to perceive configurations that are familiar to us."

Our algorithm for rhythm analysis and transcription will exploit the above cognitive concepts. The fifth principle, to bias cognition of objects into ones that we already know, is related to learning. In fact, the majority of the research in the field of quantization and transcription is concentrated on that particular principle of recognition of learned patterns. In pattern-matching solutions (Kilian 2004), or Bayesian techniques applied by Cemgil et al. (2000) and Raphael (2001), it is the user who needs to provide at least some patterns in advance. The machine needs to be trained to recognize certain patterns.

There is one more aspect of Gestalt psychology that one should not forget. The idea that there are no isolated musical events, that the term "participation in a line" expresses a deeply rooted musical and human concept, that of the individual tone taking part in an entity that means something more or rather different from all the meanings of the individual members summed together. We would like to give here two quotes from the same chapter in Efrati (1979):

> We feel a musical line to have significance when the tones group together to form a pattern, in the sense of Gestalt psychology. In this case every tone contributes to the expressivity of the whole, whilst each gets its own meaning from its position within the group, as well as from the form of the latter (Efrati 1979, p. 74).

> According to Christian von Ehrenfels a "Gestalt" is characterised by a) its supersummation (a melody signifies more than merely the sum of its tones, a face is more expressive than the sum of its constituent parts), and b) by the possibility of transposition (Efrati 1979, p. 75).

The Gestalt principle of proximity governs also the vertical axis of polyphonic ensemble practice and the perception of musical events that are supposed to start together at a certain onset time. But when scores indicate that events start together it does not necessarily mean that the onset times need to be exactly the same for these events. Physically, those onset times might be slightly different although players and listeners perceive them as being together.

> Rasch (1988) […] made recordings of three different trio ensembles (string, reed, recorder) and calculated the onset relations between tones when they were nominally simultaneous. He found that asynchrony values ranged from 30 to 50 ms, with a mean asynchrony of 36 ms. Relating these findings to his earlier perceptual ones, Rasch concluded that such onset asynchronies enabled the listener to hear the simultaneous sounds as distinct from each other. According to this line of argument, such asynchronies should not be considered as performance failures, but rather as characteristics that are useful in enabling listeners to hear concurrent voices distinctly (Deutsch 1999, p. 306).

The physical imperfection of asynchrony between voices that should start together actually facilitates polyphonic voice separation.

Another example of the strong influence of grouping mechanisms is again given by Deutsch (1999), p. 309:

> In music played by [plucked instruments], when the same tone is rapidly repeated many times, and it is periodically omitted and replaced by a different tone, the listener may perceptually generate the omitted tone. Many examples of this phenomenon occur in 20th century guitar music, such as Tarrega's *Recuerdos de Alhambra*, […] and Barrios' *Una Limosna por el Amor de Dios*. Here the strong expectations set up by the rapidly repeating notes cause the listener to "hear" these notes even when they are not being played.

Finally, music psychology research points out that timing processes are having an effect on the perception of other musical parameters and *vice versa*. In relation to the idea of inter-connectivity between single note events, research has demonstrated the importance of *temporal coherence* as a function of pitch proximity and tempo (Deutsch 1999, p. 314).

Harmonic complexity exerts a pressure within the vertical frequency-domain and on the horizontal time-domain, in so far as the perception of complex chromatic harmonies needs more time so that all aspects of voice-leading and all dynamic functions of tension and resolution are perceivable by the listener. If it goes too fast we lose out too many details and the evaluation of the dynamic functions of the sounds cannot take place. This is related also to the general fact that the frequency ratios between the sounds of a chord with increased harmonic tension are more complex than within a chord where the individual sounds have simple frequency ratios, as for example in simple triads. However, a thorough discussion of the harmonic domain is beyond the scope of this book. I would like to point the reader to Thakar (1990), and to Thakar (2011), for a further discussion.

8.7 Modelling of Neural Oscillations for Musical Meter

Studies of weakly connected neural networks (WCNN) are interested in resonance phenomena between firing neurons (Hoppensteadt and Izhikevich 1997). WCNN are reported to be useful especially for modelling pattern memorization and recognition tasks of the human brain. They employ a network of periodically spiking neurons and connection functions. In such a network, pairs of neurons are modelled as pairs of coupled oscillators. "Interaction between each pair of oscillators averages to zero unless their frequencies are commensurable" (Hoppensteadt and Izhikevich 1997, p. IX). Commensurability is established by a pair of frequencies whose ratio a/b uses relatively small and coprime integers a, b.

Large (2008) proposes a nonlinear system as a model for the neural oscillation of excitor-inhibitor cells, which allows him to develop a resonance theory of musical rhythm. Its goal is to predict the main psychological attributes of pulse and meter, which are not recognized as stimulus properties but rather as "endogenous dynamic temporal referents that shape experiences of musical rhythms".

Despite expressive tempo changes during a musical performance, metric pulsation is a musically stable phenomenon "in the sense that it can continue in the absence of a stimulus, and it possesses a form of temporal, or phase stability such that it normally synchronizes with events, but can persist in the face of rhythmic conflict, or syncopation" (Large 2008, p. 9).

Pulses can exhibit harmonic (2:1, 3:1, etc.) or subharmonic structures (1:2, 3:1), which align with the rhythmic and metric structures of a musical performance (Lerdahl and Jackendoff 1996; London 2012).

"The neural resonance theory of pulse and meter holds that listeners experience dynamic temporal patterns, and hear musical events in relation to these patterns, because they are intrinsic to the physics of the neural systems involved in perceiving, attending and responding to auditory stimuli" (Large 2008, p.29)

Many experiments with artificial pulse sequences have shown that human motor behaviour can synchronize with periodic stimuli and adapt to phase and tempo deviations (Large and Palmer 2002; Repp 2002a, 2003). Moreover, stimulus events in periodic sequences can be anticipated (Repp 2005). Pulse synchronization is also maintained in the presence of syncopation (anti-synchrony), and against more complex rhythmic information within the stimulus. Large (2008) proposes a canonical model of the actual physics of neural oscillation. The following differential equation (8.1) of the complex variable z with respect to time is derived from the Wilson–Cowan model (Large 2008, p. 19).

$$\frac{dz}{dt} = z(\alpha + i\omega + (\beta + i\delta)|z|^2) + cs(t) + h.o.t. \tag{8.1}$$

with α as the bifurcation parameter, β is the nonlinear saturation parameter, ω is the natural frequency of the oscillator, $\omega = 2\pi f$ and f in Hz, and with δ as the frequency detuning parameter. The connection strength c controls the influence of the

time-varying stimulus $s(t)$ on the oscillator. Higher order terms, $h.o.t.$, are ignored here for simplicity of the representation. This model can also be written in polar coordinates using $z = re^{i\phi}$ and Euler's formula $e^{i\phi} = \cos\phi + i\sin\phi$:

$$\begin{aligned} \frac{dr}{dt} &= r(\alpha + \beta r^2) + cs(t)\cos\phi + h.o.t. \\ \frac{d\phi}{dt} &= \omega + \delta r^2 - c\frac{s(t)}{r}\sin\phi + h.o.t. \end{aligned} \tag{8.2}$$

It is reported that the above canonical model captures certain universal properties of many known models of neural oscillations.

The mathematical analysis of the physics of neural oscillation makes several predictions in line with the resonance theory of pulse and meter. Endogenous periodicity is predicted by the spontaneous oscillations in the neural system. The generalized synchrony of pulse and meter with musical rhythms is predicted by the process of entrainment of nonlinear oscillations to an external stimulus. Finally, the perception of metrical accent is predicted by higher order resonances in nonlinear oscillators (Large 2008).

When applying the above model of neural oscillation to the analysis of performed musical rhythms, the following questions arise: When does a given pattern of note onsets fall under the categories of pulse or meter, and when does the system recognize that a pattern falls outside of those categories? There is a theoretical possibility of using chaotic oscillations and/or burst oscillations in order to understand the perception and generation of complex (non-periodic) rhythmic patterns in music (Large 2008).

However, in order to match the time scales of musical pulse and meter, it is not sufficient to use models based on single-neuron activity or alternating excitatory and inhibitory neural activities. There is evidence that burst oscillation fits indeed well with musical time scales and this line of research has received much attention (Large 2008; Large et al. 2010, 2015). As an important expansion of the neural resonance theory of musical pulse and meter, Large et al. (2010) derived a canonical model for gradient frequency neural networks (GFNN) responding to time-varying external stimuli. The new experimental outcomes, see also Angelis et al. (2013), give supporting evidence for the proposed resonance theory.

8.8 Bayesian Techniques for Meter Detection

Temperley (2007) specifically addresses the problem of automated meter detection. He cites Cemgil's and Raphael's works on quantization and tempo tracking, as described by Cemgil et al. (2000), and Raphael (2001). Temperley states that both models implicitly find the correct meter by quantization, they do not model metrical grids explicitly. Cemgil's model is reported to be capable of dealing with simple rhythmic structures from performances. Cemgil's and Raphael's models involve a

large number of parameters. As one can see from Temperley's own description, his method also works on a large number of parameters modelled by probabilistic tables that need to be learned from performances and score-representations of the Essen corpus of folksongs (Schaffrath 1995). Also, the folksong melodies in that collection clearly show only simple rhythmic structures. Temperley says that "the best way of modelling rhythm perception probabilistically remains to be determined, and may well combine ideas from all of these models. (p. 31)." His method tries to infer the metrical grid from the performed pattern of onsets according to Bayes' theorem:

$$P(grid|onsetpattern) = P(onsetpattern|grid)P(grid) \tag{8.3}$$

He uses a generative model of meter perception to calculate the most probable metrical grid to match the observed rhythm. At the lowest level, his model has a fixed grid resolution of 50 ms called "pips". The system tries to map a three-level grid based on units of pips to the performance data: A tactus level #2 is tracking the beats, a sub-tactus level #1 tracks the downbeats, and the highest level #3 models first subdivisions of the beat. The random variables needed for the modelling are derived from the Essen corpus of folksongs. A tactus interval distribution is based on experiments in music psychology (Parncutt 1994) that a stable tactus interval is approx. centred around 700 ms.[4] Temperley is aware that calculating the most probable grid on the basis on all possible grids would take a huge amount of processing time because of the number of possible grids and parameter combinations:

$$\begin{aligned} P(onsetpattern|grid)P(grid) = {} & P(UT) \times P(LT) \times P(UPh) \times P(T_1) \\ & \times \prod_{2}^{t} P(A_n) \times \prod_{2}^{t-1} P(T_n|T_{n-1}) \times \prod_{1}^{t-1} P(DB_n) \times \prod_{1}^{q} P(N_p) \end{aligned} \tag{8.4}$$

where $P(UT)$ is the probability of either a duple or triple upper level (subdivisions of the tactus), $P(LT)$is the probability of either a duple or triple lower level (downbeats), $P(UPh)$ is the probability of the phase of the upper level in relation to the tactus level (beats). $P(T_1)$ is the probability of the initial tactus interval indicating the initial tempo. This distribution prefers initial values of around 550 ms (109 BPM). Values above 700 ms, which is everything slower than approx. 86 BPM, have a combined probability of 0.01. The fastest possible initial tempo is modelled at approx. 133 BPM with a probability of 0.1. These tempo distributions are obviously style-dependent. $P(A_n)$ is the probability of generating another tactus beat, $P(T_n|T_{n-1})$ is the probability distribution of non-initial tactus intervals that is Gaussian-centred around 700 ms (86 BPM). Tempi faster than 133 BPM and tempi slower than 55 BPM are not possible. $P(DB_n)$ is the probability of the position of each lower-level beat (downbeat) in relation to the previous tactus beat. If the tactus beat is triple, then there are two probabilistic variables for each tactus beat instead. $P(N_p)$ represents the probabilities of note-onsets at pips in relation to the pip's beat level.

[4]The 700 ms tactus interval corresponds to approx. 86 BPM.

In light of intractable solutions to equation (8.4), a dynamic programming strategy is applied that concentrates on finding the different tactus levels in a stream of onset data mapped to the underlying grid of 50 ms pips. Temperley shows an example of a successful 4/4 meter detection of one of the Essen folksongs that begins with an upbeat. He tested his model trained with the Essen corpus with a subset of the corpus that was not involved in the training stage. The results show that his model is good at the beat-tracking level, the lowest metrical level, but less successful (at about 70%) with detecting the higher metrical levels, e.g., downbeats. A comparison with a previously developed 'Melisma' model shows that the older system is more successful than the Bayesian approach, with an overall score of 0.793 for the new system compared to 0.865 for the 'Melisma' model. The reason for its better performance is likely to be the consideration of actual note lengths for beat detection.

One problem reported by Temperley (2007) relates to assuming a wrong meter, or detecting the right meter but finding the beats in the wrong places (phase alignment). It is a natural argument that incorporating other musical rhythms, e.g., on harmonic or dynamic levels, might improve a system for rhythm detection that relies solely on note onset data. Yet, taking the note duration into account as for example in Raphael (2001), would require "a fundamental reworking of the system" (Temperley 2007, p. 47).

8.9 Quantization

Quantization is the mapping of performed note onset times and their durations to the onsets and durations defined by the score in common practice notation, or defined by learned metrical patterns.

8.9.1 Grid Quantization

The task of quantization, as understood traditionally in the context of sequencers and notation software, is quite simple. An arbitrary stream of performed note onsets and end-of-note onsets is placed over a predefined equidistant grid of points on a timeline. This grid is defined, for example, by the length of a 16th note, given a certain tempo in BPM. Other grid resolutions can be set by the user, for example 16th triplets or 32nd note lengths. Onsets are then placed on the arithmetically nearest point on that grid. There are various problems that arise from this strategy. First of all there is the fact that the grid has no built-in knowledge about the compound rhythms arising from using multiple subdivisions at the same time (polyrhythm). Even in its simplest form, i.e., 3 over 2, this knowledge is not represented appropriately by an equidistant grid. As a consequence, the proper differentiation between the beat division of 3 against 4 is often not possible. With a static grid quantization, tempo changes during a performance are not tracked, therefore the quantized and transcribed result of an

expressive performance is soon out of synchronization with the beat structure of the piece. Or, it can happen that durations transcribed incorrectly, which leads to an unnecessary complexity of the transcribed results - the transcribed score might get cluttered with slurs between single notes and smallest note durations connected to long ones. In order to gain some precision with grid quantization the user of a sequencer can record the performance together with a metronome click. But since real musical performances are not played against a metronome this method is of very limited use and expressive timings are ruled out from such a performance. It also happens that users might fall in front or behind the metronome click which again deteriorates the transcribed result. Because compound IOIs are not analyzed, the grid quantization method can lead to false note overlaps, i.e., note events are placed together as a chord in spite of having been played one after the other during the performance.

8.9.2 Context-Free Grammar

Longuet-Higgins and Lee (1982) proposed a method that infers a context-free grammar from a performed series of inter-onset intervals by first building IOI ratios, initializing beat duration based upon the first IOI ratios and by applying a small set of rules to keep track of changing tempi in comparison with the current tempo estimate. The rules are Initialize, Stretch, Update, Conflate and Confirm. A review of this method in comparison with other rule-based approaches can be found in Desain and Honing (1999).

Context-free grammar and their application to music analysis was highly popular during the 1980s, due to Noam Chomsky's work on the structure of the grammar of language, see also Nierhaus (2010). His main influence is evident in the GTTM by Lerdahl and Jackendoff (1996) and it has been also popularized through Leonard Bernstein's TV production of his Norton Lectures (Bernstein 1976).

8.9.3 Pattern-Based Quantization

Kilian (2004) uses a pattern-matching approach for the mapping of inter-onset intervals that were extracted from performances to score positions in Western notation. In addition, a quantization process takes into account the local context of IOIs that works on single voice streams including chords but without overlapping notes, i.e., compound rhythms that include chord onsets. He proposes a model that determines different levels of attraction between an un-quantized onset and surrounding grid time positions. Kilian states that it is not always the closest grid position that forms the correct quantized time position within the score. Therefore he applies another weighting strategy that is built from the number of common time positions between different levels of subdivision of the metrical grid, for example $\frac{1}{2}$ coincides with $\frac{2}{4}$, $\frac{4}{8}$, etcetera. In order to account for groups of onsets for which no pattern can be

found in the database, those gaps are recognized automatically and a history-based multi-grid approach is applied on their content. The histogram informs a weighting of simple IOIs based on previous IOI mappings (see also the tempo detection method described in Sect. 8.10.3 on p. 141). The examples of a pattern database provided in Kilian's Appendix contain exact score solutions for entire bars and a few more different patterns. This obviously requires a preselection from the manifold of compound rhythms that can occur within one bar. Kilian does not give further details on how to build the pattern databases in advance. The user is enabled to change the database content at any time. It suggests that the database needs to be adapted for different pieces of music and for different styles of music as well. There are no measurements provided as to what amount of information is needed by the program in advance in relation to the original score in order to perform well. Kilian also states that his model is an implicit Hidden Markov Model when he refers to research carried out by Takeda et al. (2003) who worked with learned patterns to infer score time positions and durations from performances. The pattern database for the quantization task needs to include more complex patterns as compared to the database used for tempo detection.

Kilian presents successful quantizations of a Bach Minuet, first 16 bars, including a problem with transcribing an ornament. He also shows his system's output of the first 16 bars from Beethoven's Piano Sonata Nr.20, the first movement of op. 42, 2. The meter is $\frac{4}{4}$ and contains binary and ternary pulsations (quavers and quaver triplets). In addition, he provides a transcription of the melody part of Dave Brubeck's *Take Five* with correct swing rhythm transcription.

8.9.4 Models Using Bayesian Statistics

Raphael (2001) and Cemgil et al. (2000) report systems for note onset quantization based on Bayesian inference. A performance rendering of a musical score delivers note onsets y whose positions on the time-line show a characteristic deviation from the score onsets x. Those deviations are due to a set of hidden variables θ, which can be trained with machine-learning techniques and therefore y can be traced back to the most likely score representation:

$$p(x, \theta|y) = \frac{1}{p(y)} p(y|\theta, x) p(x, \theta) \tag{8.5}$$

Their systems need to be trained with measured and analyzed data from previous performances of the score that render a probability distribution of hidden variables θ. In addition, they need to be trained with a probability distribution of onset occurrences within the score itself, i.e., when certain onsets are more likely to match with specific locations on a pre-defined metrical grid. This is dependent on style and genre and can vary even between pieces by the same composer or performer. The training stage needs to learn hidden parameters θ from previous performances that together with

the score onsets prior and the observed onsets from the actual performance lead to an optimum score representation, i.e., as close as possible to the original score. It is from this learning process of the hidden variables where some difficulties of the system arise. There are integrations of the hidden variables required that are intractable in most cases. Therefore it is required to learn the best parameter settings from given observations, which is formulated as a maximum a-posteriori estimate (Cemgil 2004):

$$\theta^* = \underset{x}{\operatorname{argmax}} \int dx p(x, \theta|y) \tag{8.6}$$

$$p(x|y) \approx p(x, \theta^*|y) \tag{8.7}$$

There are other areas as well in Cemgil's framework, where he solves intractable situations by approximation strategies.

8.9.5 IRCAM's KANT

The ICMC paper by Agon et al. (1994) describes KANT, an offline system for quantification of onsets and a library for the Common-Lisp based software PatchWork. This program, which today comes under the name OpenMusic, is a member of the family of Computer-Assisted Composition programs (CAC), where users can generated CPN scores via application of various algorithmic strategies, as exemplified in the systems by Barlow (1984), Essl (1996) or Xenakis (1992). KANT is primarily designed to facilitate the transcription of algorithmically generated rhythms and metrical structures. KANT is based on segmenting onset streams into measures, followed by an approximate calculation of the underlying beat-length and, finally, quantization of the onsets according to a user-supplied subdivision-grid mapped to the calculated beats. Segmentation is a semi-automated process via low-pass filter smoothing of the input function of discrete event durations and subsequent placement of "archi-measure" bar lines according to the second derivative of the smoothed function of event durations. Further segmentations of the archi-measures are based upon an event-density calculation where maxima are being used to trigger a further subdivision into bars. The authors say that these segmentation processes still need a certain amount of hand-editing of bar-line positions in order to achieve good results. The quantization stage works per beat of every measure and uses a least distance error strategy in order to move the original onset(s) to a position on a subdivision grid. Possible subdivisions s of the beat are in $[1 \ldots 32]$, with $s \in \mathbb{N}$. The set of quantized onsets is pre-ordered according to the least number of onsets eliminated. One of the problems with grid quantization is that separately played onsets in the original stream can be forced to share the same position on the grid, also known as false-overlaps. Through the ordering of the results, KANT favours quantizations with the least number of onset eliminations. False overlaps are generally avoided by

resolving the issue of eliminated onsets through the use of grace notes. Then, after the first ordering, two different cost functions are being applied in order to create two versions of the list of quantized solutions within the frame of one beat. The first version is ordered according to the smallest distance given by the sum of all cubic ratios of the original durations over the quantized durations. The second list is ordered by an unspecified complexity measure for the chosen beat subdivision. The authors only state the order of complexity given by the first eight possible subdivisions:

1, 2, 4, 3, 6, 8, 5, 7, etc.

That particular order of integers came from discussions with one of the authors, Joshua Fineberg. According to personal communication with Gérard Assayag, the idea is that rhythmic complexity does not follow the natural order of integers. The subdivision of a beat by relatively high prime numbers produces a more complicated score. Other contributing factors to this kind of complexity are the use of nested subdivisions and the subsequent possibility of compound ratios by using slurs between two different note durations. A beat subdivision by 1, 2 or 4 is simpler than a division by 3, because in an otherwise binary rhythm context, a triplet will be used to give a particular effect, for example a ritardando, or the triplets employed in Jazz swing. Beat subdivisions 5 and 7 are even more uncommon. A different ordering of this list might be used if the music is structured by ternary meters and rhythms. The binary subdivision by 8 has a high complexity rank because the authors do not want to get a transcription with many 32nd notes unless necessary. Because of a small onset deviation, KANT could propose a complicated transcription with many linked 32nd notes, whereas KANT would produce a very neat 16th note transcription by ignoring this small deviation,[5] hence the higher rank of the subdivision by 8.

The first solutions from both lists are then weighted by a user-supplied "precision" factor, which multiplies a Euclidean distance measure that compares both solutions with the original onset stream. After the weighting, the solution with the least Euclidean distance is chosen for rendering. KANT treats all onsets falling into inter-beat intervals in this way.

The authors do not employ any beat or downbeat detection algorithms but rather rely on onset density distributions in order to group onsets into bars of unequal lengths - a strategy that serves certain aesthetics of composition and notation in 20th and 21st century classical music. The approximative *gcd* algorithm is borrowed from solving the problem of detecting a fundamental frequency in audio signal analysis. The lengths of the measures are the input frequencies with the calculated beat-length as their fundamental. It follows that although bar-lengths are allowed to change, the beat-length is still considered to be constant and thus the underlying tempo is not allowed to change significantly. Expressive timing is only taken into account during the quantization at beat-level. KANT requires a considerable amount of user intervention at every stage of its processing, including the need to edit archi-measures and bar-lines and the need to make decisions about the weights of cost-functions employed at the beat-level quantization.

[5] According to Assayag via personal communication with the author in February 2010.

8.10 Tempo Tracking

For an excellent overview over the current state-of-the-art, I would like to refer the reader to Hainsworth (2006), which includes detailed comparisons of the various methods used for beat tracking and the analysis of musical meter.

Beat and tempo tracking are methods to measure instantaneous tempo in a musical performance on the condition that the music has an underlying pulsation. Musical performances vary through expressive timing, stylistic conventions and musical necessity. Beat tracking is a special case of tempo detection combined with score following insofar as its output is a stream of pulses coinciding with the inferred metrical structure on the highest level, whereas tempo tracking does not return the beat structure directly but through the process of quantization. Any successful quantization of onset data has to track the tempo changes involved in an expressive performance. Brute-force quantization algorithms on the other hand, like for example simple grid-quantization to the nearest 16th note, are not able to take those tempo deviations into account and they are therefore generally unsuccessful when it comes to human performances.

8.10.1 Multi-agent Systems

The system presented by Dixon (2001) is capable of estimating the tempo and position of beats by processing MIDI and audio files. The input can follow an arbitrary musical style but the music is required to have a continuous underlying beat. Dixon's method takes the duration as well as the intensity of IOIs into account. Events falling under a 70 ms threshold are merged into a single event. The number of merged onsets adds then an additional weight to the event. From this 'cleared' list of events Dixon clusters IOIs together using a 25 ms tolerance for an IOI's distance from the mean of the cluster. In addition, the method also clusters together compound IOIs. The clusters can merge if their mean values become close enough during tracking. The means are allowed to shift during the performance. If a new event lies within 25 ms distance from the mean, the respective cluster accepts this event, which in turn increases its score. Clusters whose means are integer multiples of each other receive an additional weight $f(d)$ according to the integer factor d:

$$f(d) = \begin{cases} 6-d, & 1 \leqslant d \leqslant 4 \\ 1, & 5 \leqslant d \leqslant 8 \\ 0, & otherwise \end{cases}$$

The clusters with the highest overall score are used for establishing a current tempo hypothesis. Each hypothesis is then given to an agent. The states of the agents are then continuously checked to see if they would describe a viable tempo prediction. An event may fall outside the inner window of an agent into an outer window, in

which case the current tempo hypothesis is updated according to the distance of the new event from the currently estimated beat position.

8.10.2 Probabilistic Methods

Documentation of the research in this direction can be found in Cemgil et al. (2000), Cemgil (2004), Cemgil and Kappen (2003). In these publications, it is reported that both, tempo detection and quantization can be formulated within one and the same model described as a maximum a posteriori (MAP) state estimation task. Sequential Monte-Carlo integration methods are used for finding solutions.

Bayesian statistical methods seek to infer a structure, e.g., the score in Western notation, from a known surface, e.g., a performance of a score, by application of Bayes' Rule (Temperley 2007), i.e., if the probabilities of all possible performances of a score are known with regard to the probability of the score, and if the overall probabilities of the performances and the score are known, then the following relationship can be calculated:

$$P(score|performance) \propto P(performance|score)P(score) \qquad (8.8)$$

The expression on the right side needs to be maximized in order to find the most likely score given a performance. In order to obtain the prior probability $P(score)$, the algorithm assigns statistical weights to all beat onset positions of the score on a normalized metrical grid. This requires the analysis of a known score. In order to achieve a broader application of the score prior, usually a corpus of scores from within the same style is used. It appears that such a beat tracker's success would depend on the prior distribution functions that have been chosen from specific musical styles and specific pieces. Cemgil (2004) argues that the non-predictability of a performer's decision calls for a probabilistic approach.

Raphael reports a probabilistic model for tempo detection combined with quantization methods for transcription of performed note onsets. Raphael's method splits up into a rhythm process, a tempo process involving a random walk model and the observation data from the performance. MAP estimates are being solved by using dynamic programming. As in other Bayesian techniques, the system needs to be trained with multiple score performances in the same musical style before an actual performance is being analyzed and transcribed by the system.

8.10.3 Pattern Matching

Kilian (2004), who works purely on MIDI data, uses a combination of pattern matching of groups of notes and statistical analysis for tempo detection evaluating single note durations. The pattern matching works on comparisons with a user-defined

database of short rhythmic patterns. First, the MIDI performance data is merged into a single track of onset data with added information about the number of onsets at a given point, for example chords are detected and their number of notes kept together with their approximate mean onset time, similar to Dixon (2001), see also Sect. 8.10.1 on p. 139. The MIDI velocity of the notes is also taken into account. This information informs a weight associated with each note/chord onset. Then, inter-onset interval ratios are calculated in order to detect durational accents, i.e., changes between long and short notes and vice versa. Inverse durational accents (short-long syncopations) are filtered out from the list of onsets, which leads in effect to a rule-based low-pass filtering of inter-onset interval (IOI) data. An explicit score database incorporates knowledge about rhythmic patterns in advance of the analysis. A distance function measures the distance between fragments of the performed rhythms and matched patterns of the database. The advantage of the pattern-matching approach is that the user can define the patterns within the database and expand or reduce its content. The disadvantage is that the database needs to cover all short patterns that could possibly occur in the score. There is a manifold of patterns possible even if the method is restricted to patterns within the range of the length of one bar, see for example Fig. 7.6 on Greek Verse Rhythms on p. 102. Different musical styles and differences within the same musical piece pose a challenge to the task of defining the best collection of patterns for the database. Similar to Bayesian machine learning methods, the user needs to enter information in advance of the process in order to allow the program to perform a detection task. But, patterns are allowed to overlap in Kilian's approach. A histogram-based method is used for all onsets that cannot be explained by the pattern matching algorithm. Information about score durations are being mapped to previous IOIs. This feedback information influences the weight of a zone of attraction around simpler IOI ratios. In addition, an error detection penalizes unusual score mappings within a given metrical hierarchy. Player characteristics can be taken into account by evaluating timing deviations from the IOI ratios $\frac{1}{1}$, $\frac{1}{2}$ and $\frac{2}{1}$.

8.11 Conclusion

We have described various forms of representations for musical rhythms and meter, for which Western notation uses small-integer ratios in relation to a common reference duration 1/1, the whole note. But there is, apart from the occasional use of metronome figures (BPM), no physical time incorporated in rhythm notation. In addition, apart from including words like *rubato* in the score, there exists no encoding of expressive timing deviations in the musical score, yet the flexible bending of performed durations is part of a musical performance, and it is perhaps one of the most important information channels for musicians to communicate with their audiences.

This means that an automated quantization and transcription system has a major challenge, namely to arrive at a filtering method for expressive timing. It is crucial to be able to follow unpredictable tempo changes that have musical reasons from the

point-of-view of the musician, however, it is difficult to detect them on the basis of the performance data only.

Therefore, there is a need to understand how various parts of the musical system of performance and perception works. It is crucial to understand the field temporal perception and we could learn about a range of perceptual timing thresholds. We have looked at musical meter and covered a seminal text by London (2012), in which he describes the perceptual and structural constraints for musical meter by introducing well-formedness constraints (WFC). He then also arrives at a many-meters hypothesis (MMH): Human listeners have learned throughout their life various tempo-modifications, rubati, and expressive timings, that are associated with a particular metrical pattern. This learning process is of course embedded within a specific musical culture.

Musical ornaments, like trills and arpeggios, are challenging for an automated quantization and transcription program. When executing ornaments, musicians will use a lot of freedom from metrical playing and even elements of improvisation, which can make their timing patterns unpredictable.

Throughout the next chapter, we will use the adjacent interval spectrum by Kjell Gustafson (Toussaint 2004) in order to illustrate a viable method for the detection of musical durations that have a small but continuous variation in length. It was useful to learn about Gestalt principles, which has led us to develop a grouping algorithm that is used for the classification of performed durations, and which we are going to describe in Chap. 9. For example, it is through the Gestalt principles of proximity and similarity within an adjacent interval spectrum of note durations that we are able to detect duration classes. The Gestalt principle of good continuation means that different duration classes can alternate in the course of a piece but it is still possible to differentiate between them and to detect reappearances of duration classes despite alternations and local tempo variations. Common fate is a Gestalt principle whereby groups of different duration classes are subject to the same tendency of expressive timing, i.e., a rallentando or accelerando will change the absolute time of every element of all duration classes in the same direction.

We have also discussed various methods of generating data sets, like onset detection and extraction from audio recordings, manual tapping with computer keyboard and external controllers. The noise levels vary, and lowest latencies can be achieved by using sensors connected to micro-controllers. When properly used, even computer keyboard tapping has its benefits, but one has to be careful to set audio systems operations at low latencies. Provided this is the case, errors introduced by manual tapping methods stay below the levels where they would influence the results of a quantization program.

For the analysis of musical meter we have looked at various computational models of neural oscillations (Large and Kolen 1995; Large 2008; Large et al. 2010). This is a promising line of research and it supports our argument that human listeners approximate perceived rhythms and metrical patterns by tuning in on resonant oscillatory interactions of firing neurons, which have peak resonances at small-integer ratios.

We have also reviewed recent systems of quantization and tempo tracking, which feature a range of different approaches. For quantization, these include simple grid quantizers, context-free grammars, pattern-based approaches supported by user-created databases of rhythmic patterns, algorithmic approaches and Bayesian machine learning. Tempo tracking systems use clustering in combination with a multi-agent system, Bayesian machine learning techniques, pattern-matching, and computational models of neural oscillations, which resonate to rhythmic and metrical patterns of onsets.

References

Agon, C., Assayag, G., Fineberg, J., & Rueda, C. (1994). Kant: A critique of pure quantification. In *Proceedings of the International Computer Music Conference, Aarhus, Denmark* (pp. 52–59).

al-Dīn al Urmawī, S. (1980). *Kitāb al-Adwār [Book of cycles]*. Baghdad. (al-Rajab (Ed.)).

Angelis, V., Holland, S., Upton, P. J., & Clayton, M. (2013). Testing a computational model of rhythm perception using polyrhythmic stimuli. *Journal of New Music Research*, *42*(1), 47–60. https://doi.org/10.1080/09298215.2012.718791.

Arom, S. (1991). *African polyphony and polyrhythm*. Cambridge: Cambridge University Press. (ISBN 052124160X).

Barlow, C. (1984). *Bus journey to parametron (all about çoğliuotobüsişletmesi)* (2nd ed., Vol. 21–23). Cologne: Feedback Verlag.

Barlow, C. (1991). Musiquantenlehre. Cologne. (course materials obtained by the author).

Bernstein, L. (1976). *The unanswered question*. Cambridge: Harvard University Press. (ISBN 0674920015).

Boenn, G. (2011). Automated analysis and transcription of rhythm data and their use for composition (Unpublished doctoral dissertation). University of Bath.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, *6*, 145–238.

Brossier, P. (2006). Automatic annotation of musical audio for interactive applications (Doctoral dissertation, Centre for Digital Music Queen Mary, University of London). http://aubio.org/phd/thesis/brossier06thesis.pdf. Retrieved March 2018.

Brown, J. C., & Smaragdis, P. (2004). Independent component analysis for automatic note extraction from musical trills. *Journal of the Acoustical Society of America*. http://www.merl.com/reports/docs/TR2004-078.pdf. Retrieved March 2018.

Cemgil, A. T. (2004). Bayesian music transcription (Doctoral dissertation, Radboud University of Nijmegen). http://repository.ubn.ru.nl/handle/2066/59219#.Wr5_kK3MxE5. Retrieved March 2018.

Cemgil, A. T., & Kappen, H. J. (2003). Monte Carlo methods for tempo tracking and rhythm quantization. *Journal of Artificial Intelligence Research*, *18*, 45–81.

Cemgil, A. T., Desain, P., & Kappen, H. J. (2000). Rhythm quantization for transcription. *Computer Music Journal*, *24*(2), 60–76.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.). Amsterdam: Elsevier. (ISBN 9780122135651).

Collins, N. M. (2006). Towards autonomous agents for live computer music: Realtime machine listening and interactive music systems (Doctoral dissertation, University of Cambridge). https://community.dur.ac.uk/nick.collins/thesis.html. Retrieved March 2018.

Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition*. Oxford: Oxford University Press. (ISBN 0198522576).

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, *28*(1), 29–42.
Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*(3), 341–365.
Deutsch, D. (1999). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.). Amsterdam: Elsevier. (ISBN 9780122135651).
Dixon, S. (2001). Automatic extraction of tempo and beat from expressive performances. *Journal of New Music Research*, *30*(1), 39–58.
Efrati, R. R. (1979). *The interpretation of the sonatas and partitas for solo violin and the suites for solo cello*. Zürich: Atlantis. (ISBN 9783761105504).
Essl, K. (1996). Strukturgeneratoren. Algorithmische Komposition in Echzeit. Graz: Institut für Elektronische Musik (IEM) an der Universität für Musik und darstellende Kunst in Graz. (Höldrich, R. (Ed.)).
Fraisse, P. (1963). *Psychology of time psychology of time*. New York: Harper. (ISBN 0837185564).
Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, *35*, 1–36.
Fraisse, P. (1999). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.). Amsterdam: Elsevier. (ISBN 9780122135651).
Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, ischronous sequence. *Journal of the Acoustical Society of America*, *98*(5), 2524–2531.
Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz: Evidence for a common rhythmic pattern. *Music Perception*, *19*(3), 333–349.
Gordon, J. W. (1984). Perception of attack transients in musical tones. Technical Report STAN-M-17, Stanford University, Department of Music, California, USA.
Grey, J. M. (1975). An exploration of musical timbre. Technical Report STAN-M-2, Stanford University, Department of Music, California, USA.
Grosche, P., & Müller, M. (2011). Extracting predominant local pulse information from music recordings. *IEEE Transactions on Audio, Speech, and Language Processing*, *19*(6), 1688–1701.
Hainsworth, S. (2006). Beat tracking and musical metre analysis. In A. Klapuri & M. Davy (Eds.), *Signal processing methods for music transcription*. Berlin: Springer. (ISBN 978-387-30667-4).
Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of Amercia*, *31*(6), 759–767.
Hirsh, I. J., Monohan, C. B., et al. (1990). Studies in auditory timing: 1. *Simple patterns. Perception and Psychophysics*, *47*(3), 215–226.
Hoppensteadt, F. C., & Izhikevich, E. M. (1997). *Weakly connected neural networks* (Vol. 126). New York: Springer. (ISBN 0-387-94948-8).
Huron, D. (2006). *Sweet anticipation. Music and the psychology of expectation*. Cambridge: MIT Press. (ISBN 0262083450).
Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of timbre. *Journal of the Acoustical Society of America*, *94*(5), 2595–2603.
James, W. (1890, 1950). *The principles of psychology*. New York: Dover. (Reprint. ISBN 0486203816).
Kilian, J. (2004). Inferring score level musical information from low-level musical data (Doctoral dissertation, Technische Universität Darmstadt). http://tuprints.ulb.tu-darmstadt.de/529/. Retrieved March 2018.
Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York, London: Mentor. (ASIN B0007DEOOS).
Large, E. W. (2001). Periodicity, pattern formation and metric structure. *Journal of New Music Research*.
Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *The psychology of time*. West Yorkshire: Emerald. (ISBN 9780080469775).
Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., & Kolen, J. F. (1995). Resonance and the perception of musical meter. In N. Griffith & P. M. Todd (Eds.), *Musical networks. Parallel distributed perception and performance* (pp. 65–96). Cambridge, Massachusetts: MIT Press. (ISBN 0262071819).

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.

Large, E. W., Almonte, F. V., & Velasco, M. J. (2010). A canonical model for gradient frequency neural networks. *Physica D*, *239*(12), 905–911. https://doi.org/10.1016/j.physd.2009.11.015.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, *9*(159), 1–14. https://doi.org/10.3389/fnsys.2015.00159.

Lerdahl, F., & Jackendoff, R. (1996). *A generative theory of tonal music* (Reprint of 1983 ed.). Cambridge, Mass: MIT Press. (ISBN 9780262621076).

London, J. (2009). Hearing rhythmic gestures: Moving bodies and embodied minds. World Wide Web electronic publication. http://www.people.carleton.edu/~jlondon. Retrieved February 2010.

London, J. (2012). *Hearing in time. Psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, *11*(2), 115–128.

Martínez, J. M. (2005). MPEG-7 overview [web page]. ISO/IEC. https://mpeg.chiariglione.org/standards/mpeg-7/audio. Retrieved March 2018.

Massaro, D. W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, *83*(1 pt. 1), 32–39.

Michon, J. A. (1964). Studies on subjective duration. I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica*, *22*, 441–450.

Michon, J. A. (1978). The making of the present: A tutorial review. In J. Raquin (Ed.), *Attention and performance* (Vol. VII). Hillsdale: Lawrence Erlbaum. http://www.jamichon.nl/jam_writings/1976_making_present.pdf. Retrieved March 2018.

Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Berlin: Springer. (ISBN 978-3319219448).

Nierhaus, G. (2010). *Algorithmic composition: Paradigms of automated music generation*. Wien: Springer. (ISBN 978-3-211-99915-8).

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464.

Peeters, G. (2004). A large set of audio features for sound description. Technical report. Paris: IRCAM. http://recherche.ircam.fr/equipes/analyse-synthese/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf. Retrieved March 2018.

Pöppel, E. (1972). *The study of time* (pp. 219–241). Berlin: Springer.

Raphael, C. (2001). Automated rhythm transcription. In *Proceedings of the International Symposium on Music Information Retrieval, Bloomington, IN, USA* (pp. 99–107). http://ismir2001.ismir.net/papers.html.

Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 71–90). New York, Oxford, U.K: Clarendon Press. (ISBN 978-0198508465).

Repp, B. H. (1987). The sound of two hands clapping: An exploratory study. *Journal of the Acoustical Society of America*, *81*, 1100–1109.

Repp, B. H. (2002a). Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science*, *21*(1), 1–37.

Repp, B. H. (2002b). Rate limits in sensorimotor synchronization with auditory and visual sequences. In *Meeting on Auditory Perception, Cognition, and Action, Kansas City, MO*.

Repp, B. H. (2003). Phase attraction in sensorimotor synchronization with auditory sequences: Effects of single and periodic distractors on synchronization accuracy. *Journal of Experimental Psychology-Human Perception and Performance*, *29*(2), 290–309.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*(6), 969–992.

Repp, B. H., Windsor, W. L., et al. (2002). Effects of tempo on the timing of simple musical rhythms. *Music Perception*, *19*(4), 565–597.

Right, O. (2001). *The new Grove dictionary of music and musicians* (2nd ed., Vol. 1, pp. 797–809). New York: Macmillan Publishers Ltd. (S. Sadie (Ed.). ISBN 1-56159-239-0).

Roederer, J. G. (2008). *The physics and psychophysics of music: An introduction* (4th ed.). New York: Springer. (ISBN 0387094709).

Schaffrath, H. (1995). The Essen folksong collection in kern format. Computer database. Menlo Park, CA, Center for Computer Assisted Research in the Humanities. http://ota.ahds.ac.uk/headers/1038.xml. Retrieved March 2018.

Sethares, W. A. (2007). *Rhythm and transforms*. Berlin: Springer. (ISBN 9781846286391).

Stockhausen, K. (1988). *Texte zur elektronischen und instrumentalen Musik*. Cologne: DuMont Buchverlag. (ISBN 978-3000245282).

Takeda, H. , Nishimoto, Y., & Sagayama, S. (2003). Automatic transcription of multiphonic MIDI signals. In *Proceedings of the 4th International Conference on Music Information Retrieval - ISMIR 2003* (pp. 263–264).

Temperley, D. (2007). *Music and probability*. Cambridge: MIT Press. (ISBN 0262201666).

Thakar, M. (1990). *Counterpoint*. Yale University Press. (ISBN 0300046383).

Thakar, M. (2011). *Looking for the "Harp" quartet: An investigation into musical beauty*. Rochester, NY: University of Rochester Press. (ISBN 978-1-58046-346-1).

Toussaint, G. (2004). A comparison of rhythmic similarity measures (Technical report). School of Computer Science, McGill University. http://cgm.cs.mcgill.ca/~godfried/publications/similarity.pdf. SOCS-TR-2004.6. Retrieved March 2018.

Verplank, W., Mathews, M., & Shaw, R. (2000). Scanned synthesis. In *Proceedings of the International Computer Music Conference*. Berlin: ICMA. http://hdl.handle.net/2027/spo.bbp2372.2000.235. Retrieved March 2018.

Warren, R. M. (1993). Perception of acoustic sequences: Global integration versus temporal resolution. In S. McAdams & E. Bigand (Eds.), *Thinking in sound* (pp. 37–68). Oxford: Oxford University Press. (ISBN 0198522576).

Westergaard, P. (1975). *An introduction to tonal theory*. New York: W.W. Norton. (ISBN 0393093425).

Whitfield, I. C. (1983). The neural code. In E. C. Carterette & M. L. Friedman (Eds.), *Handbook of Perception* (Vol. IV (Hearing), pp. 163–183). Orlando: Academic Press. (ISBN 978-0121619046).

Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of Experimental Psychology*, *15*(4), 357–379.

Wright, M., Cassidy, R. J., & Zbyszynski, M. F. (2004). Audio and gesture latency measurements on Linux and OSX. In *Proceedings of the International Computer Music Conference* (pp. 423–429). Miami: International Computer Music Association. (ISBN 0971319227).

Wundt, W. M. (1911). *Grundzüge der physiologischen Psychologie* (6th ed.). Leipzig: Wilhelm Engelmann.

Xenakis, Y. (1992). *Formalized music: Thought and mathematics in composition* (2nd ed.). Hillsdale, NY: Pendragon Press. (ISBN 1576470792).

Chapter 9
Rhythm Quantization

9.1 Introduction

This chapter will explain the main concepts behind rhythm quantization and transcription from a recorded performance. I will present algorithms for grouping, quantization, and transcription of recorded onset data into musical notation. The algorithms have been implemented in software as part of the *chunking* package that is available online, see Boenn (2017).

The chapter divides into two halves. The first part deals with the algorithms, specifically discussing the grouping of onset times and inter-onset-intervals (IOIs), as well as the actual quantization and transcription of the data. The second part presents the experimental framework, test data and results. I demonstrate the use of the algorithms step by step and I will give examples. The tests are based on two recordings of the Aria of the Goldberg Variations by J. S. Bach. Both recordings, one from 1955, the second one from 1981, have been made by the Canadian pianist Glenn Gould.

The grouping of onset data is a very important technique. If listeners categorize musical rhythms on the basis of simple integer ratios (Desain and Honing 2003; Large 2008; Papadelis and Papanikolaou 2004), and if they are also capable of following smooth tempo variations, which means they can correlate the rhythmic content successfully to an underlying meter (London 2012), then it is reasonable to assume that there must be a kind of window following mechanism, which buffers short sections of inter-onset-times. In this model, the process of perception would happen in a window of time that would cover the present moment. It is hard to quantify the length of the window but I propose to use 5–6 s, which roughly correlates to the cycle of one calm in-breath and one calm out-breath. It also correlates with the notion of psychological present that we have discussed in Chap. 8, Sect. 8.3.4.

It is conceivable that there must be yet another process, which groups similar inter-onset-times into duration classes, whose relation to each other is represented by a quasi-simple integer ratio. When changing the tempo smoothly, those duration classes have to shift along with the direction of the change, in order to keep a sense of a stable rhythmic structure. One can regard this phenomenon as a manifestation of the Gestalt-principles of 'similarity', 'good continuation' and 'common fate', see Sect. 8.6. Rhythmic stability is then defined by relatively simple ratios between performed event onsets that group together into duration classes. The stability is maintained in the performance of those duration ratios although their absolute duration values, which can be measured in seconds, change during an accelerando or during a ritardando. Because of this, a duration class can only be meaningful when it can be related to other duration classes in the temporal vicinity with whom relatively simple integer ratios can be established. This means that although there are variations within each duration class, the mean ratio between different ratio classes should remain more or less constant, at least within the present moment. Their absolute timing values might change significantly over the course of a longer passage of music. Once we have established the ratios between duration classes, the actual task of transcription should become relatively easy.

The algorithm begins with an ordered set of onset times extracted from a musical performance. The grouping algorithm works on windowed sets of onset data that are one bar long. The classification of IOIs is carried out per individual window, where similar IOIs are clustered into groups. Finally, I will also show an algorithm, which arrives at notated representations of duration classes for each analysis window. In the second part of the chapter, the experimental framework will be presented and the test results are going to be evaluated.

9.2 Grouping of Onsets into Duration Classes

Let us analyze an ordered set of note onsets S from a musical performance. Remember that S can be recorded in various ways, see Sect. 8.4. In my experiments, the onset detection from audio recordings was done by using the *aubio* library by Brossier (2006). It is a good library, but one still has to check and edit the markers that it returns because of the possibility of false positives, in which case too many onsets were detected. Or, it could be the case that one or more onsets were not detected. The unit is seconds, and the precision of the markers is to the 6th digit beyond the decimal point, or 10^{-6} s. For the following description, assume that the set of onset times has already been edited and that it is always perfect.

$$S = \{s_1, s_2, s_3, s_i, \ldots, s_{|S|}\} \tag{9.1}$$

The objective of the quantization algorithm is to filter out expressive timing, and to subsequently transcribe S into Western notation. The process begins with the classification of IOIs into duration classes according to the following method.

Based on S, an ordered set of inter-onset intervals (IOIs) D is formed:

$$D = \{(s_2 - s_1), (s_3 - s_2), (s_{i+1} - s_i), \ldots, (s_{|S|} - s_{|S|-1})\} \tag{9.2}$$

The IOIs represent the time intervals elapsed between note events. Very often, IOIs are equivalent to the note durations written in the score. The set D is split up into n analysis windows W, where each window contains a sequence of IOIs:

$$W = \{d_1, d_2, d_i, \ldots, d_{|W|}\} \tag{9.3}$$

The length of the window W depends on the type of algorithm that has been used. I have worked with three different methods:

1. Downbeat markers that are set in D by hand.
2. An algorithmic segmentation process sets a mark, whenever it detects a change from long to short IOIs in D in line with certain conditions.
3. A sliding window analysis.

The first windowing process is the most reliable at the time of writing, and that is the method I am concentrating on in this chapter. The grouping algorithm per window continues as follows: First, the durations in W are sorted in ascending order. After the comparison of the differences between consecutive elements with a given threshold, τ seconds, the durations are ordered into k different duration classes T_k:

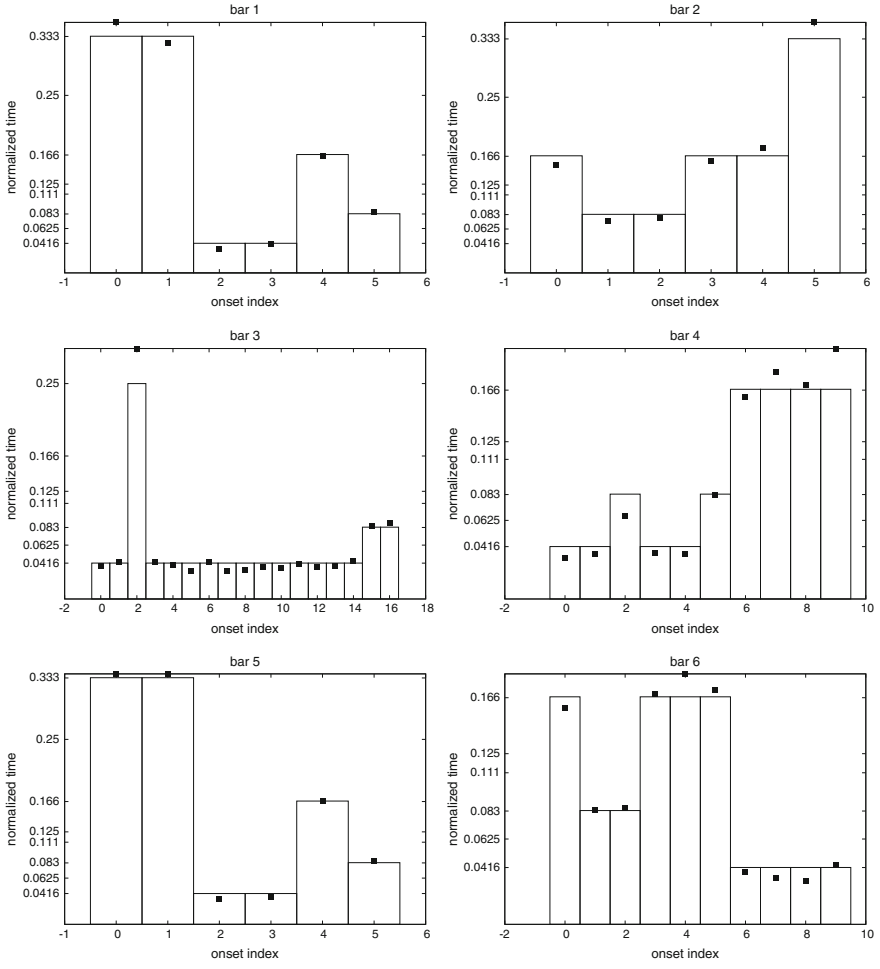
$$T_k = \{d_{i+1} | (d_{i+1} - d_i) < (\tau), i = 0, 1, 2, \ldots, |W| - 2\} \tag{9.4}$$

with $\tau = 0.1$ s, and $d_0 \in T_0$. The grouping continues until all durations have been uniquely assigned to a class T_k. The value for τ is very important. In this case, it was chosen according to the 100 ms threshold that was discussed earlier, in Sect. 8.3.2. The perception of a note as a rhythmic value within the framework of a meter, is dependent on its duration. It has to be at least tau seconds. The number of duration classes k per window can vary between 1 and $|W|$, but usually there are not that many. In the case of the Aria of the Goldberg Variations, which has many ornaments, τ can also assume values as low as 0.04 s. This is in aid of capturing the smaller durations of some of the trills that are played by Glenn Gould.

After sorting all durations in W into classes T_k, the durations are normalized and the mean for each class is calculated. The means of the classes T_k will be quantized against a set of ratios that is part of the metrical grid derived from a Farey Sequence. The sum of all durations in the analysis window is 1. Therefore, the performed durations are regarded as a normalized L1 norm vector, just like a Farey Sequence.

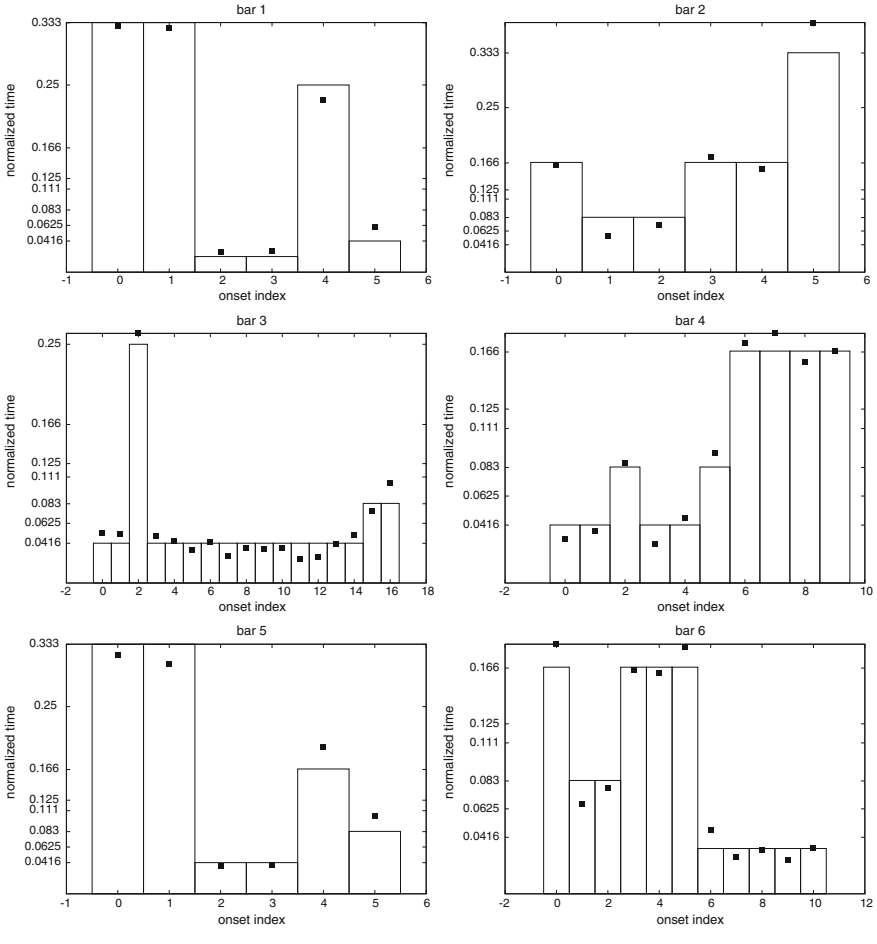
Table 9.1 shows the durations measured from the audio recording of Bach's Aria by Glenn Gould, together with the correct grouping into duration classes. The bar graphs display the quantized values that can be transcribed immediately into notation. The algorithm's work on Gould's first recording of the Aria, see Table 9.2, is

Table 9.1 Quantization of Gould's 1981 recording of the Aria, bars 1 to 6



equally successful in grouping and quantizing the onsets. One can also observe that the expressive timing is quite different between the two recordings, which is not surprising, of course. The earlier recording has also a much higher tempo, although there was no need to take the difference in tempo into account while the algorithm was applied to both sets of data. The appendix on page 175 shows the complete sets of data from both performances.

Table 9.2 Quantization of Gould's 1955 recording of the Aria, bars 1 to 6



9.3 Quantization to a Metrical Grid

One of the key ingredients of the whole process is the metrical grid against which the mean duration classes are being quantized. For this purpose, a filtered Farey Sequence is used. Recall that the mean durations and the elements of a Farey Sequence lie all in the range between 0 and 1. This range has as a reference length of one bar $= 1$, and it is independent from the tempo of the performance. Then, the quantization has to select a mapping between the mean durations and one particular element of the metrical grid. There are two filtering methods with which the algorithms have produced good results: A so-called smooth-number filter, which I described in Sect. 7.3. The other option is a filter that creates a metric grid, see Sect. 7.3.7. Table 9.3 shows an example

Table 9.3 Metric grid defining possible ratios p/q for quantizing note durations in a ternary metre. For example, if the metre is 3/4, then the duration of the full bar corresponds to one. This case is shown in the last column on the right

Ratio	Weight	p	q	Shorthand
0.016666667	11.06666667	1	60	`[10.]`
0.020833333	6.666666667	1	48	`[8.]`
0.027777778	7.333333333	1	36	`[6.]`
0.03125	5	1	32	`[16-]`
0.033333333	10.06666667	1	30	`[5.]`
0.041666667	5.666666667	1	24	`[4.]`
0.055555556	6.333333333	1	18	`[3.]`
0.0625	4	1	16	`[8-]`
0.066666667	9.066666667	1	15	`[5I]`
0.072916667	7.666666667	7	96	`[8-]~[16.]`
0.083333333	4.666666667	1	12	`[.]`
0.111111111	5.333333333	1	9	`[3I]`
0.125	3	1	8	`[4-]`
0.133333333	9.066666667	2	15	`[5H]`
0.145833333	6.666666667	7	48	`[4-]~[8.]`
0.166666667	3.666666667	1	6	`.`
0.208333333	5.666666667	5	24	`.~[4.]`
0.222222222	5.333333333	2	9	`[3H]`
0.25	2	1	4	`[-]`
0.266666667	9.066666667	4	15	`[5W]`
0.291666667	5.666666667	7	24	`[-]~[4.]`
0.333333333	2.666666667	1	3	`I`
0.416666667	4.666666667	5	12	`I~[.]`
0.444444444	5.333333333	4	9	`[3W]`
0.5	1	1	2	`-`
0.583333333	4.666666667	7	12	`I~[-]`
0.666666667	2.666666667	2	3	`H`
0.833333333	3.666666667	5	6	`H~.`
0.888888889	5.333333333	8	9	`[3W~W]`
1	0	1	1	`H~I`

of a metric grid used for the quantization of the Bach Aria recorded by Glenn Gould. In both cases, every fraction in the filtered Farey Sequence is assigned a weight according to the prime factor analysis of their denominators, which is carried out with our version of Barlow's function, see Eq. 7.8 in Sect. 7.3.6.

Within the filtered Farey Sequences, the algorithm can search either for a quantized onset position or a quantized duration. In the examples that I present here, the latter option is used. First, a search area is defined for the mean of the normalized duration classes. This area extends from 50% below the duration class to 50% above. Within this window, the algorithm searches for elements of the Farey Sequence that have the highest weight attached. The values of the weights are scaled according to the central part of a Gaussian that runs over the extent of the search area. The table for the Gaussian is filled with values of the function

$$f(x) = exp(-x^2) \tag{9.5}$$

sampled at 1024 equidistant points in the range between $x = -1$ and $x = 1$, see Fig. 9.1. The algorithm accepts the top six elements from this process into a list of candidates that represent one particular duration class. Next, from the lists of candidates, and based on the order of the duration classes in the original sequence of markers, all possible solutions are constructed. All possible combinations between the elements of the lists are being generated with no repetitions allowed. The number of combinations is equal to the number of durations within a duration class to the power of the number of duration classes. Every possible solution is then evaluated using a set of cost functions. The cost functions are:

1. The sum of Barlow's function, see Eq. 7.8, over the denominators of the duration ratios, i.e. the ratios between two consecutive note lengths. A lower value is a good indicator for a lower complexity of the notated output, which means that more complicated transitions between two durations bear a higher cost and will contribute to a lower rank of the solution.
2. The Euclidean distance between the quantized solution and the original list of durations (IOIs) that were recorded.
3. The sum of Barlow's function, see Eq. 7.8, over the numerators of the quantized durations. This function ranks those durations lower that represent dotted note values.

The Euclidean distance measure is used as follows. Out of the set of possible combinations of quantized durations, let Q be a normalized set of quantized durations:

$$Q = \{q_1, q_2, q_i, \ldots q_{|W|}\},$$

and let W be the normalized set of recorded durations:

$$W = \{d_1, d_2, d_i, \ldots, d_{|W|}\},$$

then we use the Euclidean distance measure

$$d_E(Q, W) = \sqrt{\sum_{i=1}^{|W|}(q_i - d_i)^2} \tag{9.6}$$

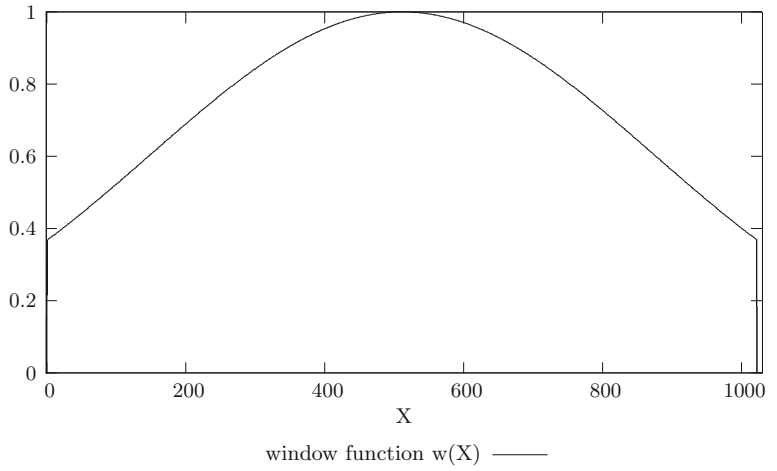


Fig. 9.1 $f(x) = exp(-x^2)$ sampled in $[-1 \ldots 1]$

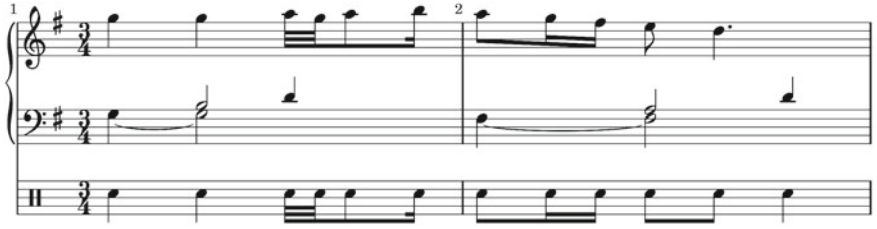



Fig. 9.2 The first 2 bars of the *Aria* of the *Goldberg Variations*. The third staff shows the compound rhythm of Gould's 1981 recording that has been generated from the means of the duration classes T_k, which were detected by the algorithm and then quantized for a score representation

With this distance measure we can find out how far a particular set of quantized durations deviates from the actual performance.

Continuing with the list of solutions, it is sorted in ascending order, first according to cost function #1 (sum of Barlow's function, Eq. 7.8, over all denominators of the duration ratios), then again sorted in ascending order according to cost function #3 (sum of Barlow's function, Eq. 7.8, over the numerators). The top ten solutions are retained for one further evaluation. They are sorted in ascending order according to cost function #2 (Euclidean distance). The first solution is selected, and/or the first solution, whose cost function #3 equals zero.

After all durations have been quantized in this way, a transcription into notation can follow. Because the reference of the analysis window is a full bar with a specific metric structure, every transcription is a simple mapping process between the quantized data and the symbols of the SNMR. Finally, the SNMR symbols can be transcribed directly into Western notation. Both steps are completed by a simple table lookup procedure (Fig. 9.2).

9.4 Some Further Examples of Grouping

The following figures show the compound rhythms of the opening bars of Herbie Hancock's piece *Chameleon* from the record *Head Hunters*. Figure 9.3 shows the rhythm of the bass-line pattern, which is quite similar to the 3-2 *son clave* pattern. A characteristic difference is the three eighth-notes anacrusis, which creates an interesting tension. The way that the bass-line is being played on a synthesizer makes the perception of the downbeat quite ambiguous. This tension lasts until the drum-set starts, see Fig. 9.5, which resolves the riddle. Of course for the listener who knows the piece very well this effect deteriorates quite a bit. Our grouping algorithm detects all duration classes of the compound rhythms successfully as shown in Figs. 9.4 and 9.6.

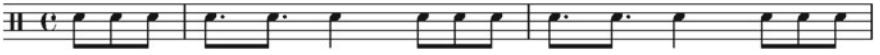



Fig. 9.3 *Chameleon*: bass-line pattern

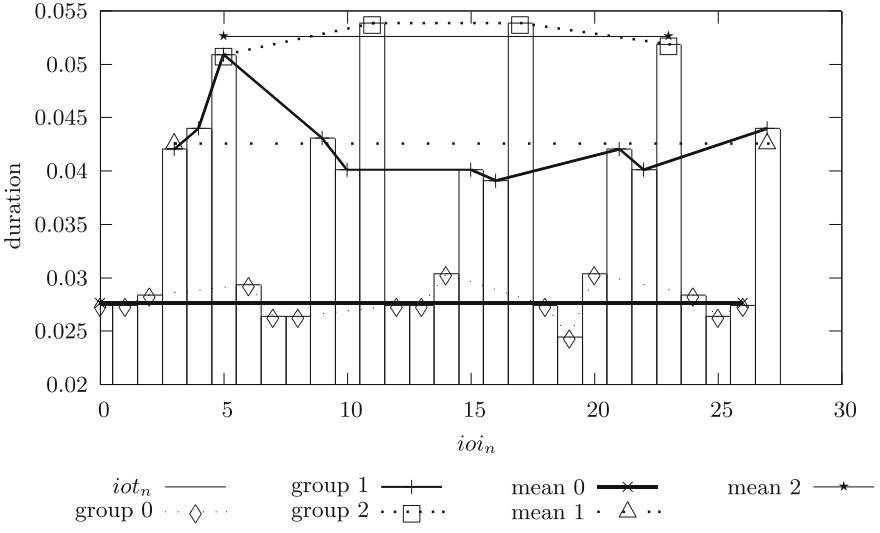


Fig. 9.4 Groups detection for Herbie Hancock's *Chameleon*: Rhythm of the bass-line at the beginning. An approximation of the mean values yields: 1/37, 1/24, 1/19, which is close to the original set of IOIs 1/36, 1/24, 1/18 or 1/18, 1/12, 1/9. Transcribed into a 4/4 bar, 1/18 represents the quaver, 1/12 a dotted quaver, and 1/9 represents the crotchet of the rhythmic pattern. Note that the variation of the IOIs strongly indicate a human player rather than a quantizing sequencer program at work

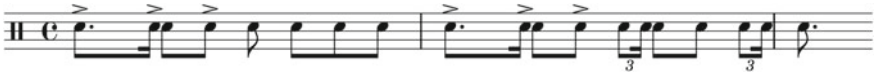



Fig. 9.5 *Chameleon*: compound bass-line and drums pattern

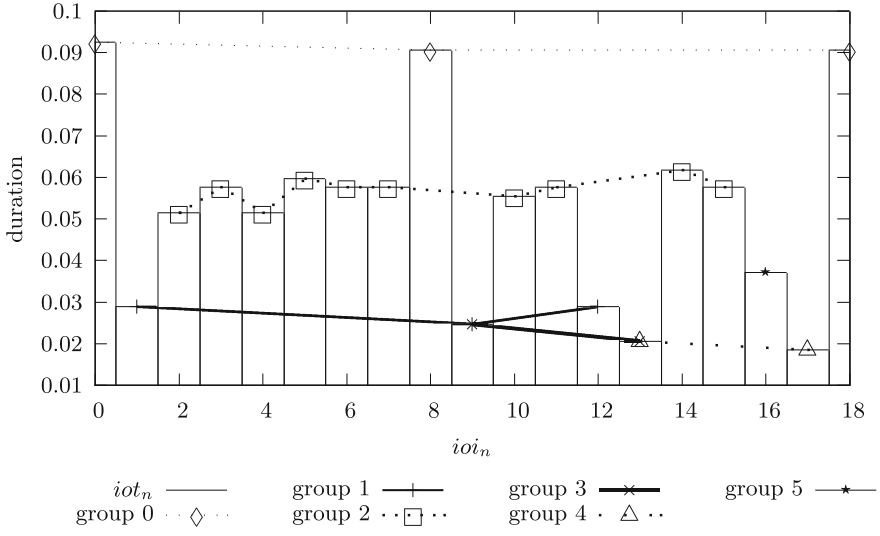


Fig. 9.6 Groups detection for Herbie Hancock's *Chameleon*: Rhythm of the bass-line now joined by the drum set

9.5 Quantization of Onsets to a Filtered Farey Sequence

Apart from the quantization of performed note durations to a metric grid, I also developed a method to quantize the onset positions within the window of analysis. In this case, a filtered Farey Sequence F_{200} represents non-equidistant grid points for quantization. The filtering process uses k-smooth numbers; in our case we tested the algorithm with 3-smooth numbers, which is sufficient for a broad number of musical cases, as we have seen in Sect. 7.3. The order 200 for the Farey Sequence grid creates a fine timing resolution for the analysis window of 5 per mil of the total length of the window. For example, it would enable a resolution of 100 ms for a 20 s analysis window. The mean values of all duration classes $\{M(T_k)\}$ per analysis window are calculated. They are then processed by the quantization algorithm. From here, the remaining steps are similar to what I have described earlier. A search window is formed around the means, followed by the collection of duration class candidates from the filtered Farey Sequences. These candidates have weights attached that are further scaled by the same Gaussian function, and so on, see the pseudocode in Algorithm 2.

Algorithm 2 Attach weights to integer ratios of filtered F_{200}

```
▷ Input: mean, range, list                                   ▷ Output: modified list
max ⇐ (mean + 0.5 × range)
min ⇐ (mean − 0.5 × range)
length ⇐ length of Gaussian table
while not past the end of list do
    element ⇐ current element of list
    if element.value ≥ min ∧ element.value ≤ max then
        n ⇐ length × (element.value − min)/(max − min)
        w ⇐ value of Gaussian table at index n
        element.weight ⇐ w/(indigestibility of element.value.denominator)
    else
        element.weight ⇐ 0
    end if
end while
```

9.6 The Transcription Algorithm

If the analysis window aligns with one full bar of note onsets, then the transcription is a simple table lookup procedure. Either an SNMR output or a LilyPond script is generated. But, if an analysis window is not aligned exactly with one bar in the music, then the following approach can be used.

9.6.1 Analysis Windows of Arbitrary Length

It starts with the sets of normalized quantized durations, one for each analysis window. Per set one calculates various metrical grids. Then, one tests how well they align with the onset points provided by the quantized durations.

The metrical grids are based on equidistant beats. The number of beats per grid are taken from this set of prime numbers:

$$B = \{2, 3, 5, 7, 11, 13, 17\}$$

In Western notation, beats have a duration that is given as an integer ratio just like note durations. We therefore form a set of integer ratios with the elements of B:

$$C = \{\frac{2}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \frac{11}{2}, \frac{13}{2}, \frac{17}{2}\}$$

As denominator one can chose a different power of 2, because in notation we also encounter the denominators 4, 8, 16 and 32. However, it is not so much a question of which denominator of the meter is chosen, i.e. the user could chose a different power of 2. What does matter is the number of beats of the underlying metrical grid that we are going to estimate here. This is encoded in the numerator of the time signature of the meter.

Let Q_i be a set of quantized and normalized durations of analysis window W:

$$Q_i = \{q_1, q_2, q_j, \ldots q_{|W|}\},$$

with $j = 1, 2, 3, \ldots, |W|$. We then calculate k scaled sets of durations per element c_k of C, with $k = 1, 2, 3, \ldots, |C|$:

$$Z_k = \{q_j \times c_k, j = 1, \ldots |W|\}$$

Then, for each of the rational numbers z_j in the set Z_k, we calculate Barlow's harmonicity function according to Eq. 7.10 and sum the results together.

$$H(Z_k) = \sum_{j=1}^{|W|} H(z_j)$$

The smallest sum of harmonicities $H(Z_k)$ marks the best fitting beat duration c_k, in the sense that a transcription of Z_k would require only very few slurs, if any, between note durations and that only those subdivisions of the beat are preferred in the transcription that have relatively low prime factors. This means that, for example, quintuplets, septuplets, 11-tuplets, etcetera, are less likely to appear in the transcription. Once the best beat duration is chosen, Z_k is fed into a subsequent transcription algorithm.

9.7 Experimental Framework

In this section, we would like to set out how we are going to test the main hypothesis of this chapter: Can filtered Farey Sequences be used to model musical rhythm and meter that have been extracted from a musical performance? As we have seen in Chap. 7, it is possible to use the method of filtered Farey Sequences together with scores, to analyze different kinds of meters, to evaluate cyclical rhythmic patterns, or to work with lists of natural integers. We have also seen that the metrical tempo grid proposed by London (2012) can be modelled algorithmically by using a filtered Farey Sequence. There is a specific interest for the quantization of performed onset times of compound rhythms. For example, a piano score involves two hands playing in polyphony. It is their combined action, the compound rhythm of their performance, which creates the flow of music that we hear. What is heard is then sent to neural processes in the brain that specialize in recognizing learned patterns, such as musical meter, and also rhythmic patterns that are composed over an underlying musical pulsation or meter. We have seen in Sect. 8.4 that the transients of sounds carry most of the crucial timing and rhythm information. This justifies the interest in the analysis and transcription of onsets recorded from performed compound rhythms.

We will now introduce our test material and propose a distance measurement, which helps to determine the quality and success of the quantization and transcription process.

9.7.1 Test Material

We have chosen two recordings of the *Aria* of the *Goldberg Variations* by J.S. Bach played by Glenn Gould in 1955 and in 1981. Onsets in both recordings were extracted by using the algorithms of the *aubio* library by Brossier (2006). A certain amount of hand editing was necessary to delete wrongly detected onsets or in order to place undetected onset markers. Both onset streams were segmented into analysis windows by hand. We have tested the quantization and transcription algorithm by using windows that were aligned to the length of one bar in $3/4-$meter.

The Bach *Aria* is rhythmically interesting because it contains many ornaments, which have been notated by Bach using special symbols as markups below or above the note, for example trills, mordants, appoggiaturas etc. This means that these parts of the piece are not written down in detail and that it is left to the performer to interpret these signs and to execute them freely, i.e. still within the metric framework, but also with expressive timing. It follows that the exact way that a certain ornament is played is not predictable, which is a useful challenge for any quantization and transcription algorithm. I have chosen Gould's performances because they are extremely different from each other, although it is the same musician performing. Gould's playing is an interesting case for the analysis of expressive timing and its temporal micro-structure (Bazzana 1997; Oswald 1997), not the least because of the conscious choices made by Gould to contrast his last recording with the early rendering from 1955. The tempo of the *Aria* is almost twice as slow in 1981. In addition to the abundance of different ornaments, the *Aria* features syncopated rhythms, room for showing expressive phrasing, all in contrast with a smooth sequence of 16th notes in the last quarter of the piece, see score in Fig. 9.13, page 167, bars 27 ff.

9.7.2 Distance Measurements

As part of the quantization process, the quantized sets of durations are analyzed in terms of their distance from the performed durations. We also included an analysis of the distance of the quantized set from the original section of score. We have evaluated the outcomes of our quantizer by using the same Euclidean distance measure as before. See Toussaint (2004) for a review of various distance mesaures in order to express similarity between different rhythmic patterns.

9.7.2.1 Euclidean Distance

Given two n-dimensional vectors of durations (IOIs),

$$X = \{x_1, x_2, x_3, \ldots, x_n\}$$

and

$$Y = \{y_1, y_2, y_3, \ldots, y_n\},$$

the Euclidean distance $d_E(X, Y)$ is measured by:

$$d_E(X, Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{9.7}$$

Clearly, when the distance between a quantized set of durations and the original set of durations within the score is equal or very close to zero, we should see a perfect or very close match between them, which is the aim of the quantization system. Having a short distance between the quantized result and the performed durations is also a useful feature of the algorithm. Some of the expressive timing can show up in the transcription, especially in the cases of ornaments, and with ritardandos at the end of sections.

9.8 Test Results

Here we present a series of test results that demonstrate the outcome of our grouping, quantizer and transcription algorithms. For this series of tests, the knowledge of the downbeats is given to the program.

The following Figs. 9.7, 9.8, 9.9, 9.10, 9.11, 9.12, 9.13 and 9.14 display the original score of the *Aria* with all of Bach's ornaments written out in the same way as they were played by Gould. The third staff, below the piano staves, shows the results of the quantizer. In this way, one can see how the quantized durations relate to the durations in Bach's score. What the quantized part shows are the compound rhythms that are produced by both hands during the performance.

When it comes to Gould's execution of Bach's ornaments, there are notable differences between the two recordings. The reader may compare the last figure in the right hand in bar 6, the trills in the left hand in bars 8, 10, 21 and 22. An appoggiatura means to execute a note a little earlier than the following onset on one of the main beats. This is what occurs in 1981 in bar 19, left hand, just before the third beat. The following trill in the central voice is also played differently. Finally, there is a note inégale in the right hand in bar 26 that occurs only in 1981. Ornamentations are always open to variations by the individual player, both with regard to the timing and with

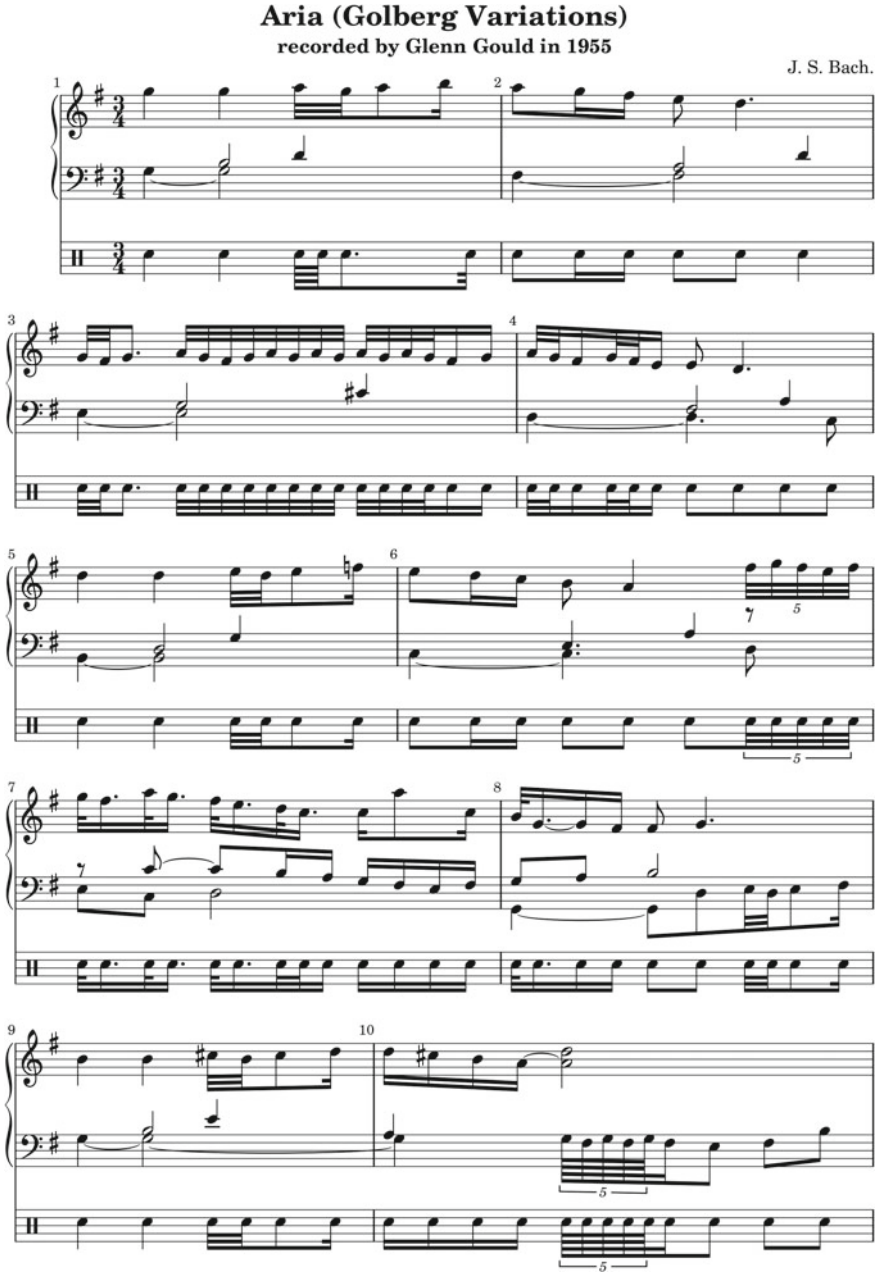



Fig. 9.7 Quantization of Gould's recording of the Bach Aria in 1955

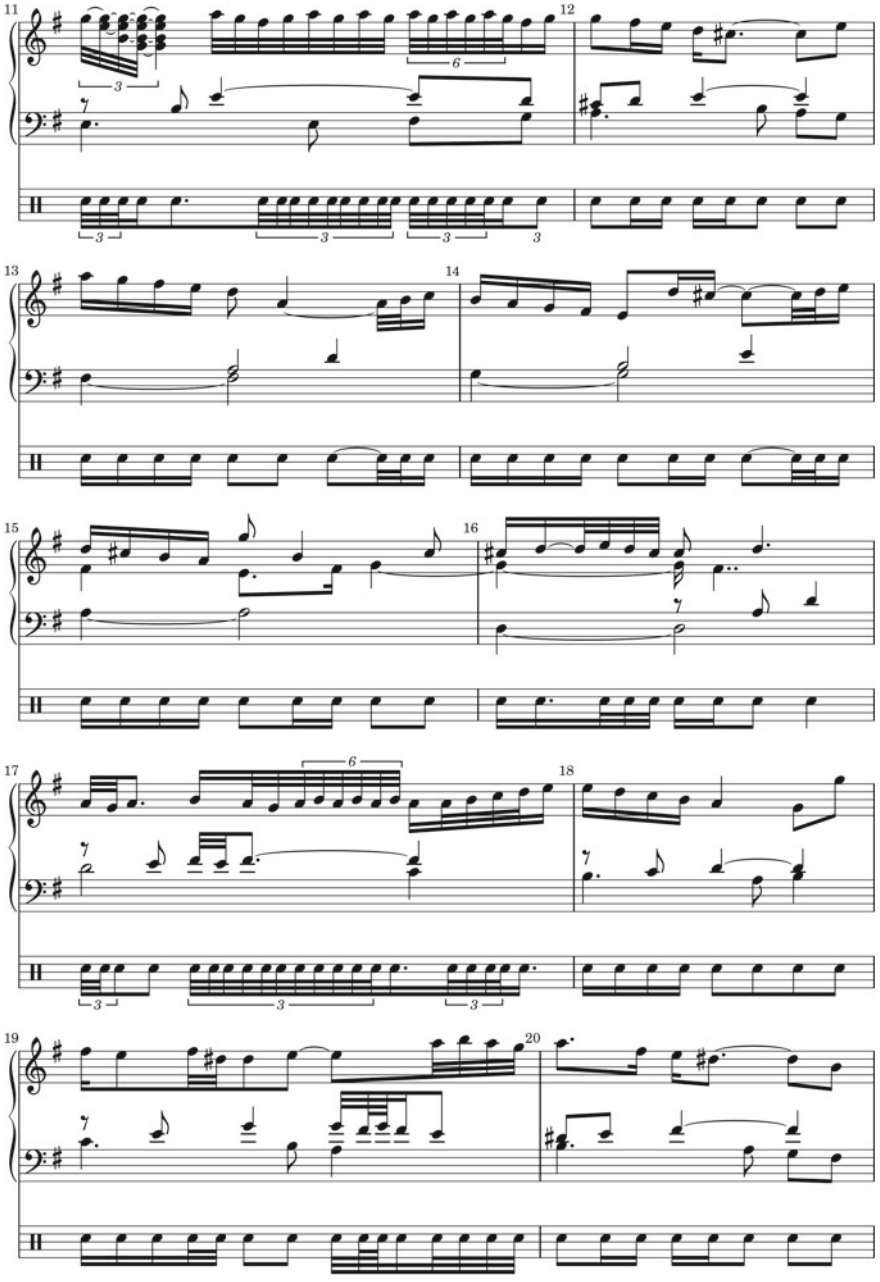



Fig. 9.8 Quantization of Gould's recording of the Bach Aria in 1955

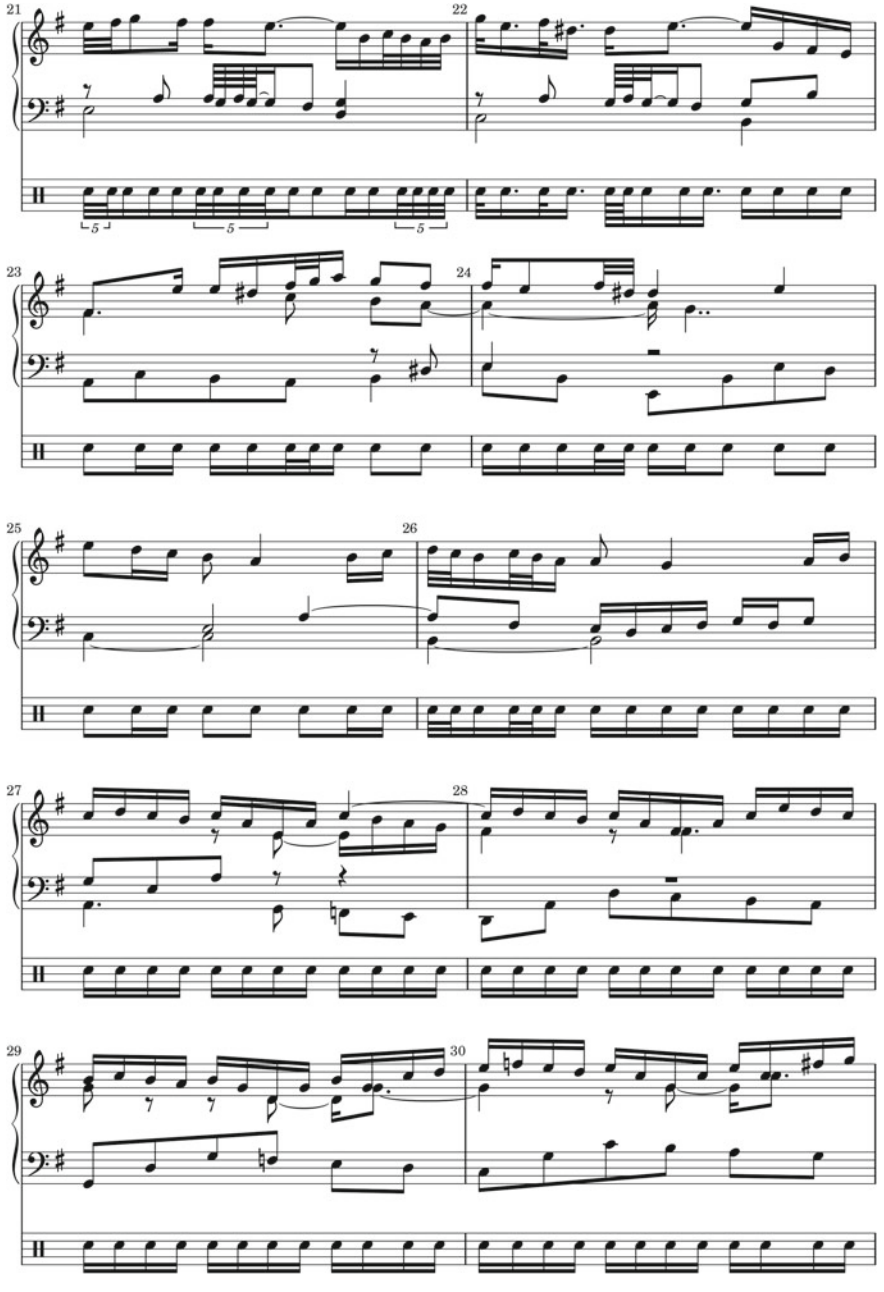



Fig. 9.9 Quantization of Gould's recording of the Bach Aria in 1955

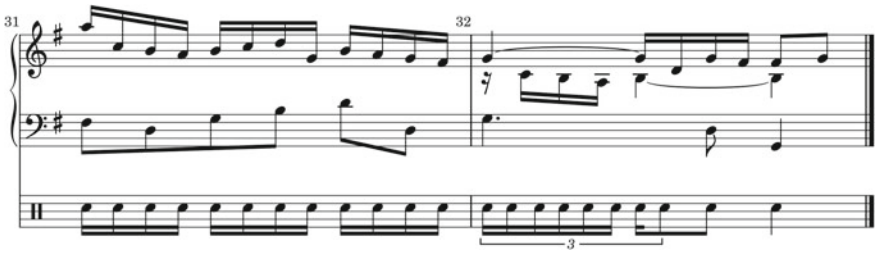



Fig. 9.10 Quantization of Gould's recording of the Bach Aria in 1955

regard to the rhythmic structure of the particular ornament. Therefore, ornaments are a particularly challenging task for any quantization and transcription program.

I also compared the performed durations per bar with the original score and calculated the Euclidean distance $d_E(W_l, J_l)$, with J_l representing the normalized set of compound note durations in bar $l = \{1, 2, 3, \ldots, 32\}$, as written in the score and including a metrical notation of the way in which Gould has played Bach's ornamentation. The largest distance one can detect between Gould's performance in 1955 and the score is in bar 22, with $d_E(W_{22}, J_{22}) = 0.207624$. In Gould's 1981 performance, the maximum Euclidean distance is in bar 21: $d_E(W_{21}, J_{21}) = 0.165846$. Tables 9.4 and 9.5 give an overview of the Euclidean distances between Gould's performance and the original score in the 4th column. The next column shows the distances between the quantized bar and the original score. Here, a zero distance indicates a perfect match between the two.

9.8.1 Observations

The 1955 recording is much faster than Gould's last recording of this work. In bar one of the 1955 recording, Gould plays the ornament on beat 3 (we call this a mordent) with 64th notes, and the last note changes into a 32nd note versus the 16th note that is in Bach's score. This is the consequence of the freedom that is granted to the player, especially when it comes to Baroque music. But, compare this also with bar one from the 1981 recording where the tempo has calmed down significantly. Here, the ornament is executed with 32nd notes followed by an eighth note and a 16th note. The quantization process has spotted this fact and it is able to correctly represent the outcome of the decisions that Gould has made. It is also interesting to look at the ornament in bar 6 on the last eighth note and compare the two recordings. In 1955, we hear a 32nd quintuplet, but in 1981 it has one note less, and the rhythm has four 32nd notes instead. Again, it is good that the program could make that distinction. In bar 10, we find again an ornament that was played differently. The left hand trill has fewer notes in 1981 compared to 1955, but this time, the subdivision in both cases is in 64th quintuplets followed by a 16th note. Bar 11 is a big challenge because of

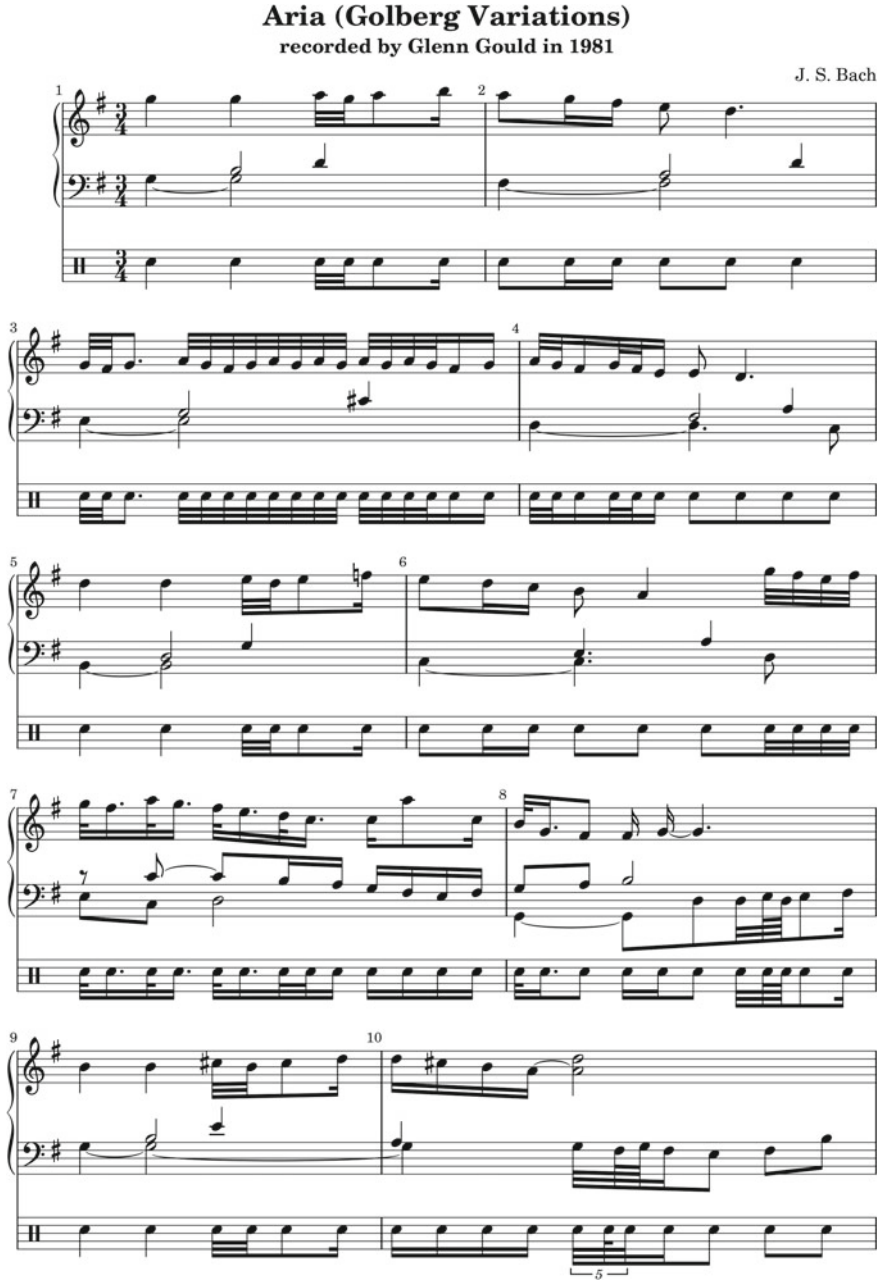



Fig. 9.11 Quantization of Gould's recording of the Bach Aria in 1981

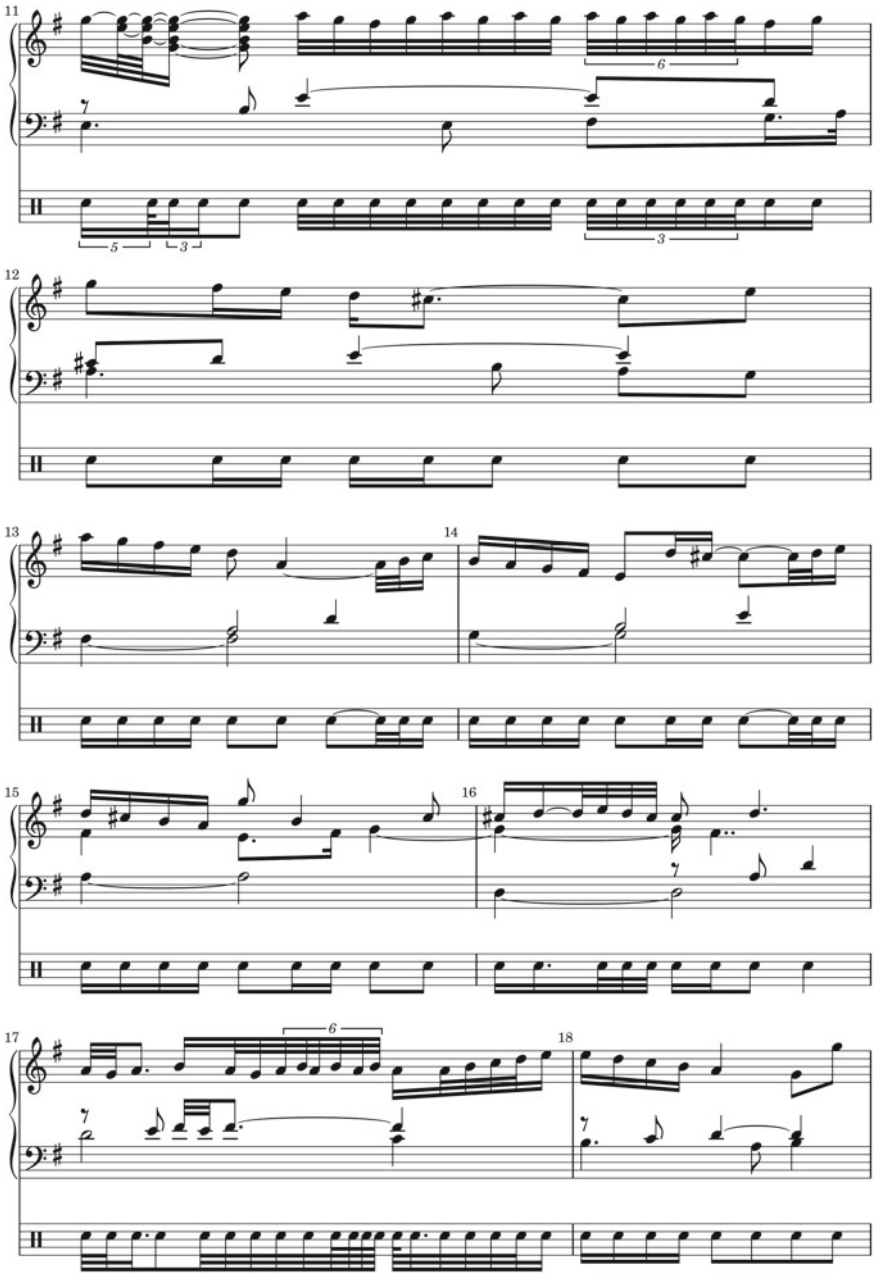



Fig. 9.12 Quantization of Gould's recording of the Bach Aria in 1981

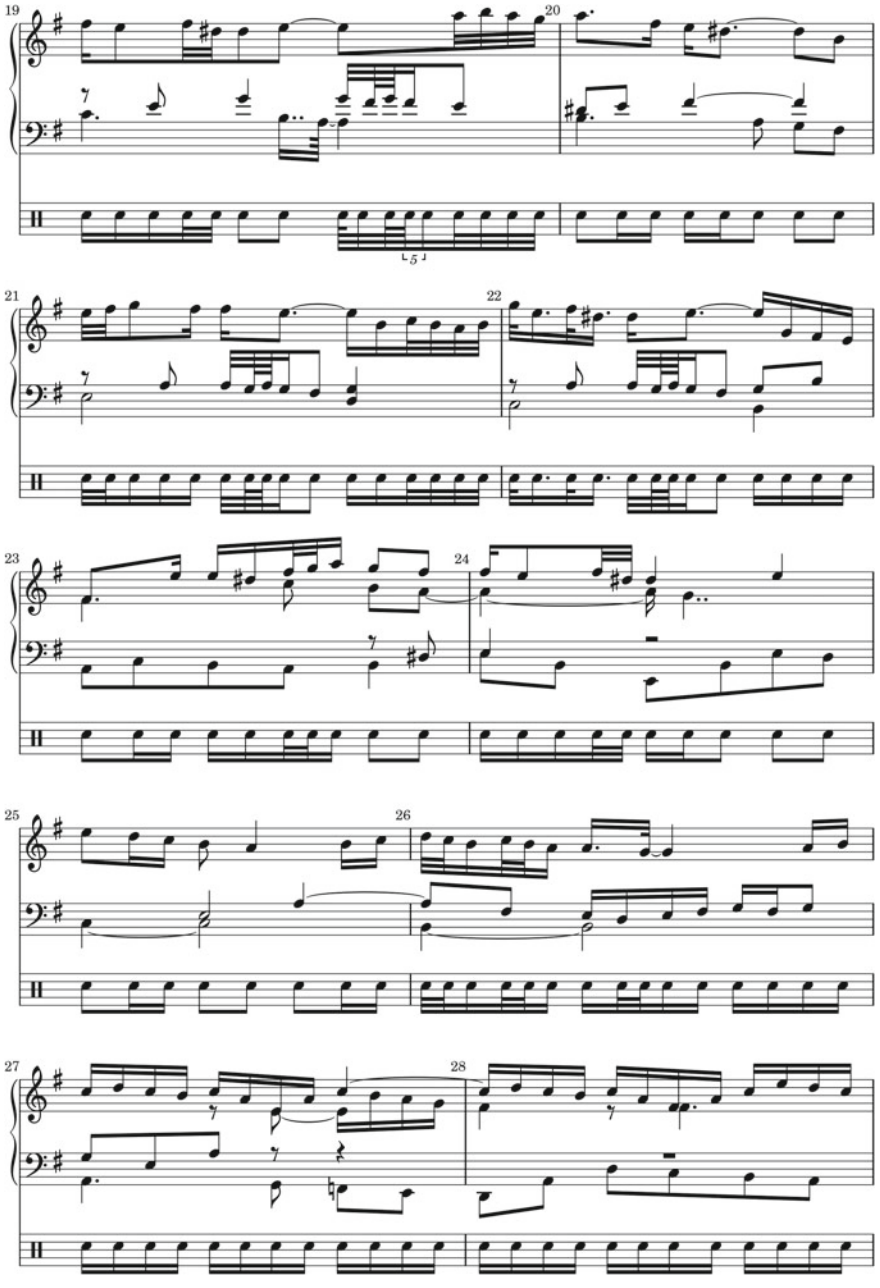



Fig. 9.13 Quantization of Gould's recording of the Bach Aria in 1981

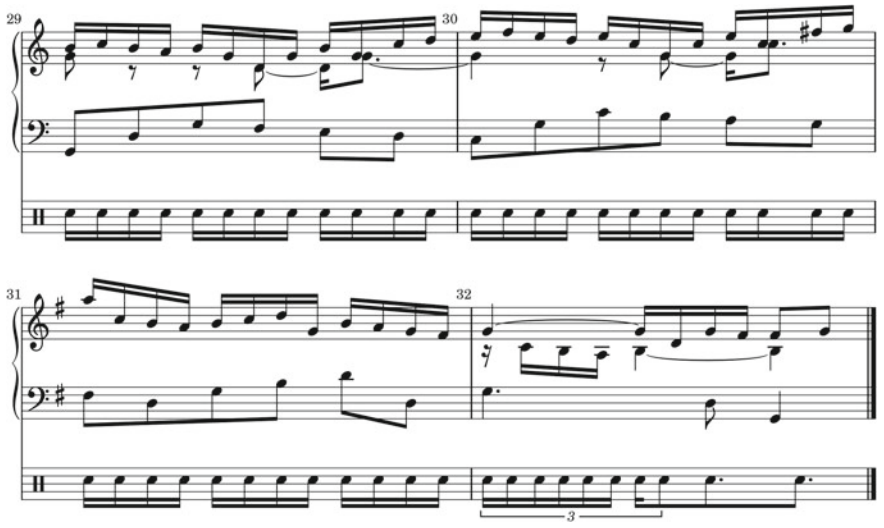



Fig. 9.14 Quantization of Gould's recording of the Bach Aria in 1981

the long trill that speeds up in both versions, and because of the downward arpeggio that is played on beat 1. In the case of the 1981 recording, we see a good match between the transcription and the quantized performance. The transcription of the 1955 version starts the trill one 16th note too late. Because of that, the first part of the trill is notated too fast. But, on the other hand, the second part of the trill matches perfectly, and the remaining durations represent a small ritardando towards the end of the bar.

Bar 17 poses another big challenge. Two ornaments, one in the left hand, and one in the right hand, are starting on the same beat (beat number 2). In 1955, Gould plays the left hand slightly ahead of the tempo, which produces what one refers to as *notes inégales*. They are common practice, for example in French Baroque music. The effect is that the differences between durations in the left and right hand become so tiny that the quantization algorithm has difficulty to integrate them into a full bar of measured durations. Therefore, the differences between the smaller notes are resolved as a collection of 32nd triplets. Note also that it is possible to distribute the triplets in unusual ways. Normally, triplets appear in groups of three, but in the transcription of bar 17 of 1955 they can come also in a group of 11, and in a group of four. The total sum, 15, is a multiple of three, and that is why there is no problem to align the durations within one bar. The algorithm always checks whether this is the case. This strategy is one way of notating *notes inégales*. The same happened to the 1981 transcription, only here it is the dotted notes that represent a slight prolongation of duration, which lead to a slight metric shift inside the metric bar structure. Something similar happens in bar 21 of the 1955 recording, and in bar 19 of the 1981 recording.

Table 9.4 Euclidean distances measured between the normalized durations of Gould's 1955 performance and Bach's score (4th column). The 5th column shows the Euclidean distance between the quantized performance and the score. See also Figs. 9.7, 9.8, 9.9 and 9.10

Analysis of the quantization of Gould's 1955 performance					
Bar	Number of onsets	Bar length in seconds	Euclidean distance betw. *Gould's performance* and original score	Euclidean distance betw. *quantized performance* and original score	Bar
1	6	2.7575	0.0674779	0	1
2	6	2.8095	0.0339418	0	2
3	17	2.4205	0.0362363	0	3
4	10	2.9295	0.026371	0	4
5	6	2.926	0.0424895	0	5
6	11	3.036	0.0341105	0	6
7	13	2.9615	0.0331713	0	7
8	10	3.872	0.0709284	0	8
9	6	3.106	0.0420389	0	9
10	13	3.207	0.0575668	0	10
11	21	2.9875	0.0230705	0.096225	11
12	8	3.329	0.0455098	0	12
13	9	3.5205	0.0213791	0	13
14	10	3.6545	0.0264608	0	14
15	9	3.7325	0.0302118	0	15
16	9	4.2545	0.034634	0	16
17	20	3.2245	0.115762	0.125	17
18	8	3.6135	0.0313371	0	18
19	15	3.4015	0.0370465	0	19
20	8	3.524	0.0226643	0	20
21	17	3.457	0.0336027	0.089365	21
22	14	3.552	0.207624	0.093169	22
23	10	3.901	0.0280434	0	23
24	10	4.069	0.022005	0	24
25	8	3.3905	0.0186024	0	25
26	14	3.401	0.0373512	0	26
27	12	3.4795	0.0160471	0	27
28	12	3.503	0.0169429	0	28
29	12	3.3875	0.0285099	0	29
30	12	3.485	0.0162152	0	30
31	12	3.835	0.0281443	0	31
32	10	6.232	0.0739067	0.0833332	32

Table 9.5 Euclidean distances measured between the normalized durations of Gould's 1981 performance and Bach's score (4th column). The 5th column shows the Euclidean distance between the quantized performance and the score. See also Figs. 9.11, 9.12, 9.13 and 9.14

Analysis of the quantization of Gould's 1981 performance					
Bar	Number of onsets	Bar length in seconds	Euclidean distance betw. *Gould's performance* and original score	Euclidean distance betw. *quantized performance* and original score	Bar
1	6	4.831	0.0243751	0	1
2	6	5.324	0.0224464	0	2
3	17	4.72	0.047737	0	3
4	10	5.584	0.0285121	0	4
5	6	5.584	0.0110546	0	5
6	10	5.503	0.0230716	0	6
7	13	5.488	0.0199194	0	7
8	11	6.177	0.0202216	0	8
9	6	5.869	0.0271574	0	9
10	11	5.875	0.0238282	0	10
11	22	5.474	0.0293574	0	11
12	8	6.169	0.020688	0	12
13	9	5.897	0.0867668	0	13
14	10	6.084	0.0867668	0	14
15	9	6.257	0.0186715	0	15
16	9	6.844	0.025393	0	16
17	20	4.876	0.0244918	0.051031	17
18	8	5.568	0.0211017	0	18
19	16	5.323	0.0957753	0.027003	19
20	8	5.475	0.013019	0	20
21	16	5.253	0.165846	0	21
22	13	5.444	0.0214225	0	22
23	10	5.491	0.0112145	0	23
24	10	6.256	0.0259337	0	24
25	8	6.293	0.022554	0	25
26	15	5.52	0.0143151	0	26
27	12	5.549	0.00669801	0	27
28	12	5.497	0.0125099	0	28
29	12	5.544	0.00899706	0	29
30	12	5.712	0.0126862	0	30
31	12	6.067	0.0150426	0	31
32	10	10.275	0.0867668	0.11453	32

The last bars in both recordings contain a final ritardando, a considerable slowing down of the tempo. Both quantized transcription reflect this by averaging out the tempo change over one bar, and then notating the ritardando out using smaller note durations that gradually get larger. One can interpret this as an attempt to represent a quantized version of Gould's expressive timing in those last bars.

9.9 Conclusion

The method that I presented in this chapter is successfully able to quantize onsets from performance data extracted from audio recordings. The resulting transcription perfectly matches the underlying score in almost all of the cases that we have tested. An additional challenge in these tests was to find a transcription for the many ornaments that Bach has put into the Aria. The execution of the ornaments gives some room for the player to improvise. This has a bearing on the number of notes that are being played, as well as it has a big influence on the the timing of the note onsets. We have seen cases where two ornaments occur at the same time, which gives rise to some metric shifts in the quantized performance and transcription of bar 17, for example. Expressive timing is more than just a deviation from the written note values. The roots for expression in time are in the ups and downs of musical energy. The ebbs and flows of this energy are felt as impulse and resolution, which themselves may come from different sources, for example from changes in harmony. In melody, impulse stems from the different scale degrees that the pitches participate in. In rhythm, we have variations of density, metric shifts, and the position of a note onset within a metric structure, each of which play their own roles in generating musical impulse. Some metric shifts are so strong, for example when a performer plays *notes inégales*, that they manifest in the quantized performance transcription as metric shifts as well.

The grouping algorithm works very well and its outcomes confirm previous results with regard to the 0.1 s threshold for those note durations that would still fit within the structure of a meter. The threshold is independent of the musical tempo, see Sects. 8.3 and 8.3.2 in particular. It is possible to tell the algorithm to group durations with respect to much smaller thresholds, down to 0.04 s. This is necessary if one wants to transcribe trills and other ornaments correctly but it can lead to slight metric shifts in the transcription of the performance data.

We saw that Barlow's function, see Eq. 7.8 on page 105, is useful for choosing the best candidates for the quantization. His harmonicity function, see Eq. 7.10 on page 106, is also useful for finding the best metric grid for an analysis window in an automated way and thereby finding the least complex notation. However, the user is able to impose a specific meter to be used for the transcription task.

The particular challenges of the Bach *Aria* are caused by its many simultaneously executed ornaments, which are occasionally combined with syncopated rhythms. In addition, we are dealing with polyphonic voices between two hands, which can produce some rather complicated compound rhythms.

9.9.1 Some Final Thoughts

How could expressive timing be analyzed and what could be learned from the analysis of specific performances? Is the difference between the notation of rhythms in a score and their manifestation in a performance leading to answers about how we engage perceptually with expressive timing and how we perform and listen to music that unfolds in time?

The research presented in this book has tried to give answers to these questions. In the future, we will continue to ask about the relation between a piece of music that is written in form of a modern score and its performance. How precisely do they differ? We might find more answers if we concentrate the effort on the elements of music that are not directly expressed in a score, for example the tempo.

The choice of a tempo is being born out of the necessity to realize the dynamic function of music, which follows the build-up and release of musical energy that is taking place through a synthesis of all musical factors at any moment of a performance and where the past events influence the present actions; where the experience of the present moment leads to expectations and causes future musical events. Clearly, these circumstances render every performance to be unique and unrepeatable.

Time does not exist independently from temporal phenomena. There is a danger of misunderstanding time and music through the practice of written-out scores in so far as a score might lead to the illusion of time as a "grand vessel" (empty staff lines) in which musical events (notes) are being placed as if an absolute time-line existed and thus the score and all its musical events would have an absolute and independent existence of their own. From this perspective, it remains to be a challenge to reverse the process of a musical performance, i.e., to start with a recorded performance and to infer a score from the data. The research presented in this last chapter showed that it is possible under the condition that the knowledge about the downbeat is given in advance. It is also important to have a clean data set of note onsets, which is an area that still requires editing. The next challenge for this research will be to develop a reliable beat tracking algorithm, and to advance the method of a sliding window analysis.

References

Bazzana, K. (1997). *Glenn Gould: The performer in the work: A study in performance practice*. Oxford: Oxford University Press. (ISBN 0198166567).

Boenn, G. (2017). *Chunking - Algorithmic Composition and Rhythm Analysis in C++*. https://github.com/gboenn/chunking.

Brossier, P. (2006). *Automatic Annotation of Musical Audio for Interactive Applications* (Doctoral dissertation, Centre for Digital Music Queen Mary, University of London). http://aubio.org/phd/thesis/brossier06thesis.pdf (Retrieved March 2018).

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*(3), 341–365.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *The psychology of time*. West Yorkshire: Emerald. (ISBN 9780080469775).

London, J. (2012). *Hearing in time psychological aspects of musical meter* (2nd ed.). Oxford: Oxford University Press. (ISBN 978-0199744374).

Oswald, P. F. (1997). *Glenn Gould: the ecstasy and tradegy of Genius*. New York: W.W. Norton & Company. (ISBN 0393318478).

Papadelis, G., & Papanikolaou, G. (2004). The perceptual space between and within musical rhythm categories. In Davidson, J. W. (Ed.), *The music practitioner: Research for the music performer, teacher, and listener* (pp. 117–129). Burlington: Ashgate. (ISBN 0754604659).

Toussaint, G. (2004). *A Comparison of Rhythmic Similarity Measures* (Technical Report). School of Computer Science, McGill University. http://cgm.cs.mcgill.ca/~godfried/publications/similarity.pdf (SOCS-TR-2004.6. Retrieved March 2018)

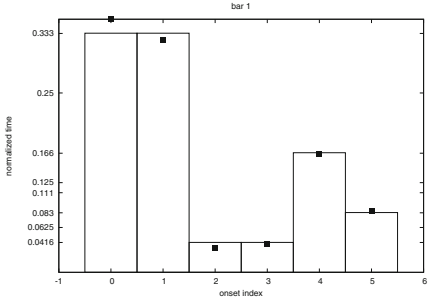
Appendix

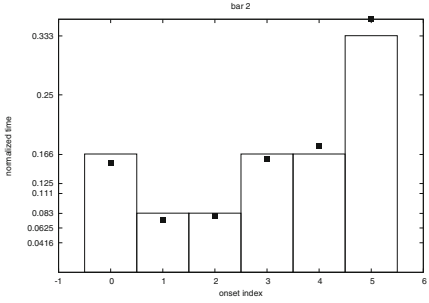
A.1 Quantization of Recorded Performances

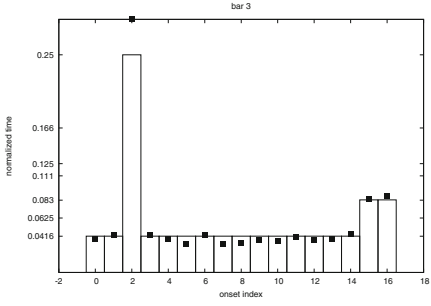
The following Tables A.1, A.2, A.3 andA.4 show the complete sets of data of Glenn Gould's 1981 recording of Bach's Aria. The bargraphs represent the quantized note durations and the successful grouping of the inter-onset intervals. The black squares denote the recorded note durations.

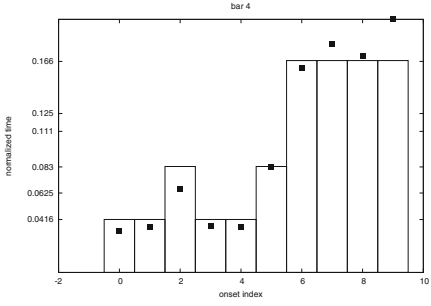
Tables A.5, A.6, A.7 and A.8 shows the same for Glenn Gould's 1955 recording of Bach's Aria. Complete sets of data are given. They can be compared to the score renderings shown in Chap. 9.

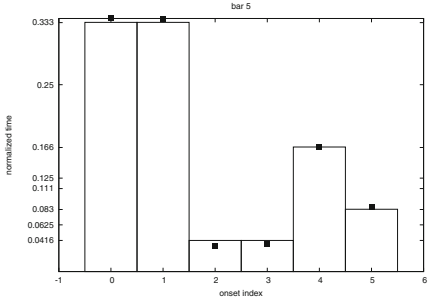
Table A.1 Quantization of Gould's 1981 recording of the Aria. Bars 1–8

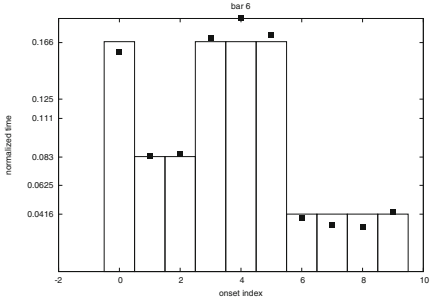


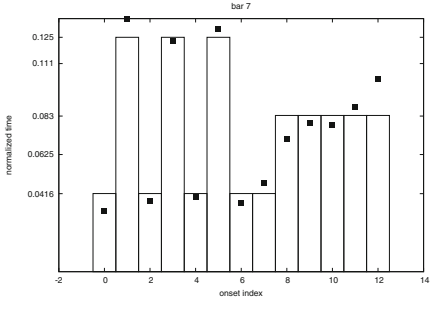












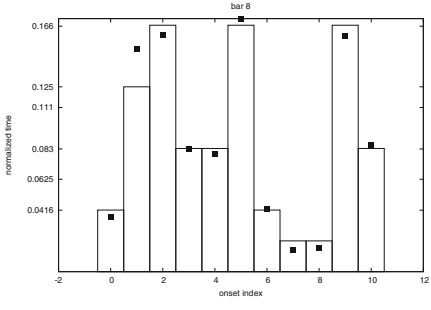
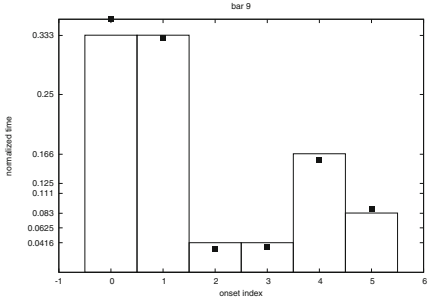
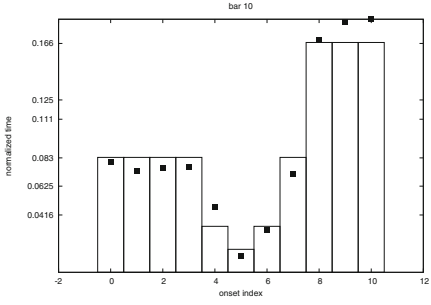
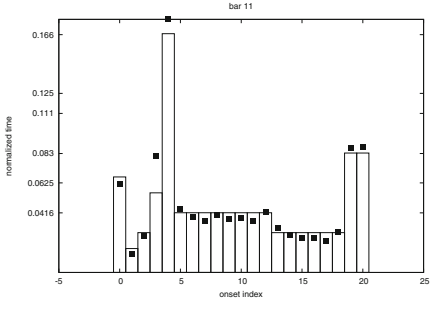
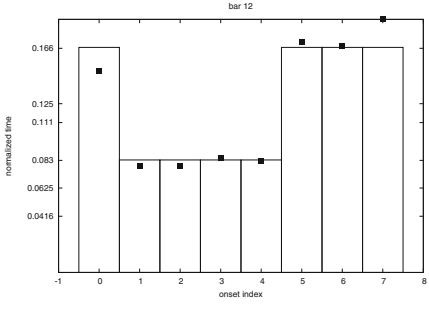


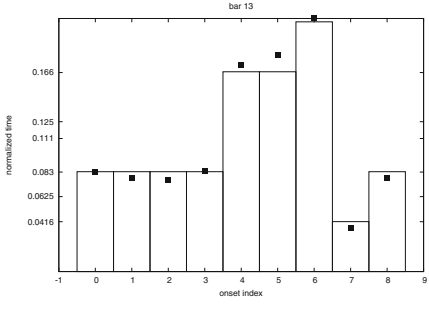
Table A.2 Quantization of Gould's 1981 recording of the Aria. Bars 9–16

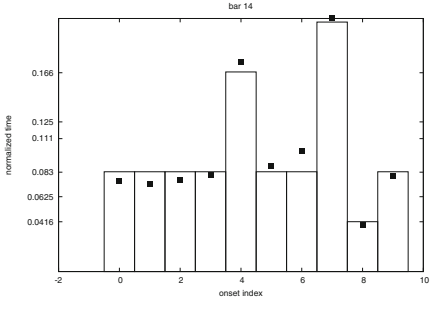


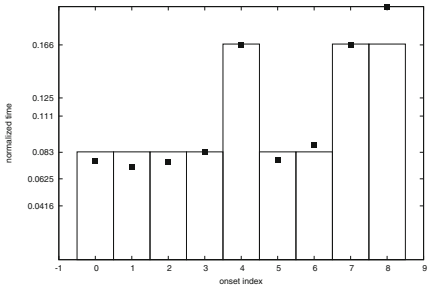












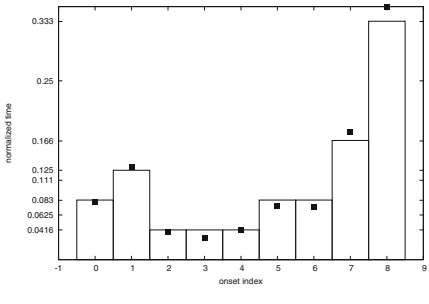


Table A.3 Quantization of Gould's 1981 recording of the Aria. Bars 17–24

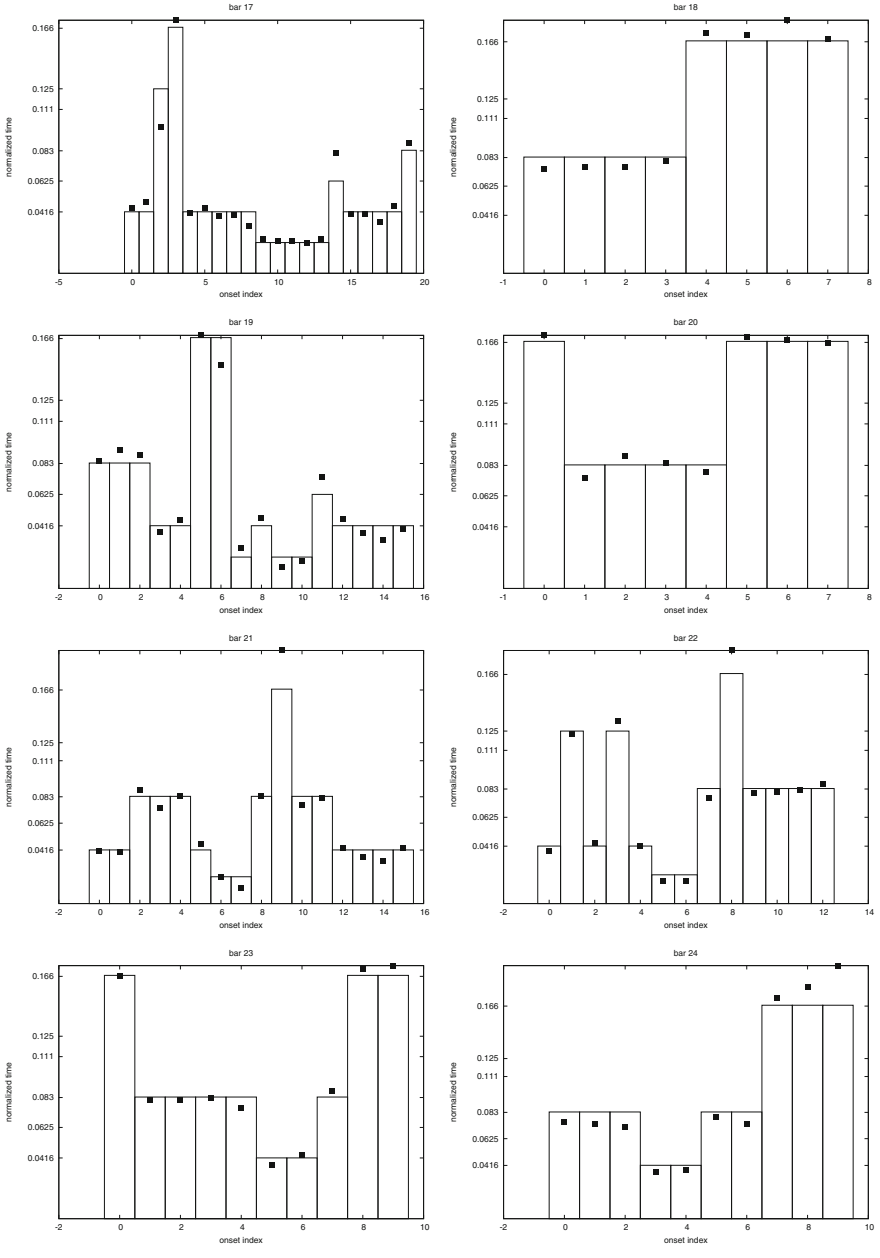
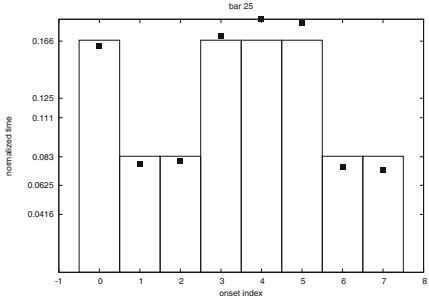
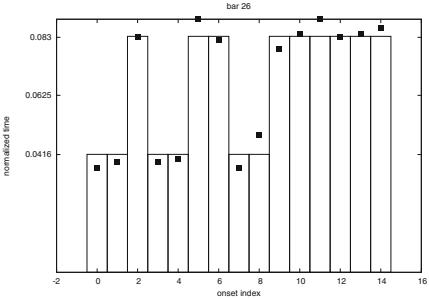
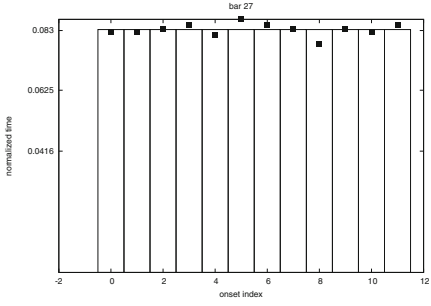
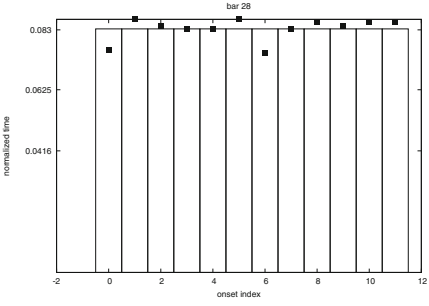


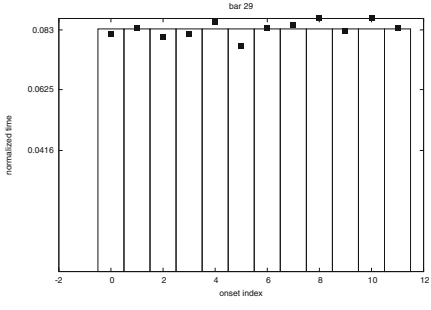
Table A.4 Quantization of Gould's 1981 recording of the Aria. Bars 25–32

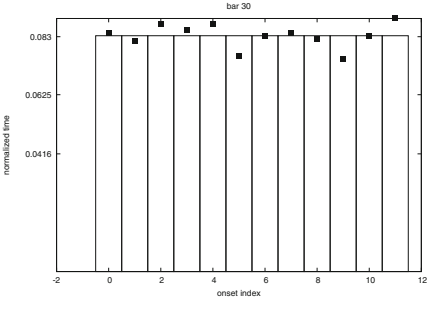


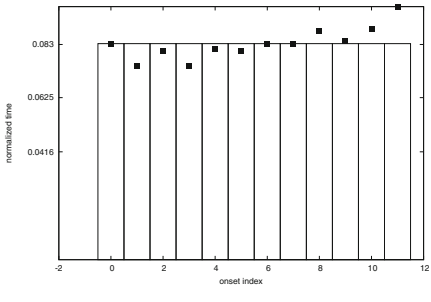












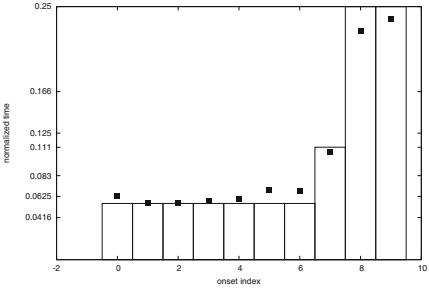
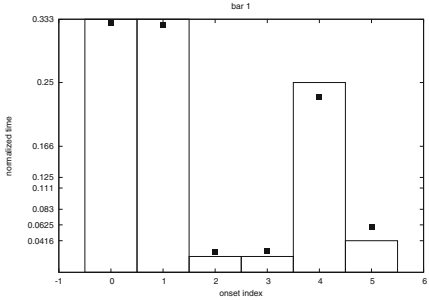
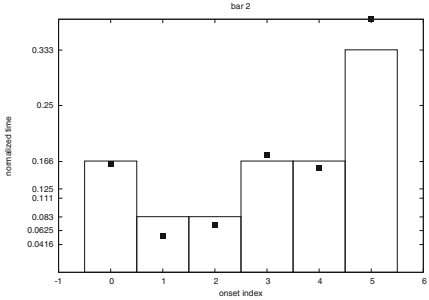
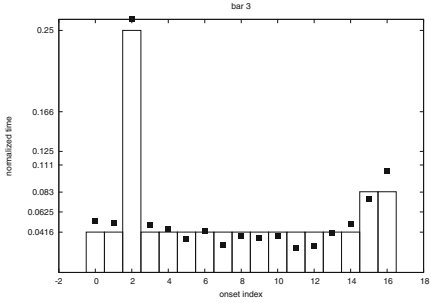
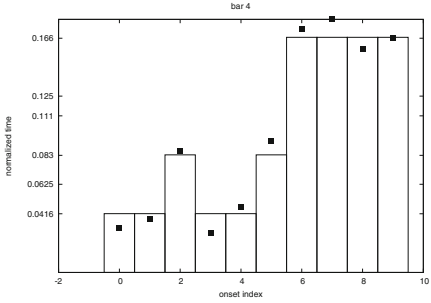


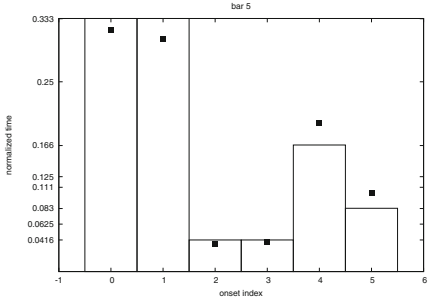
Table A.5 Quantization of Gould's 1955 recording of the Aria

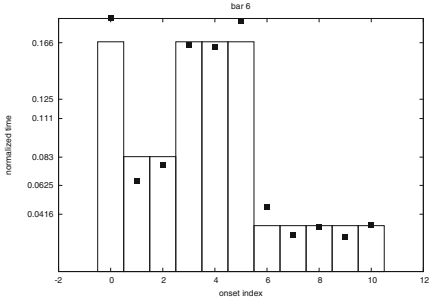


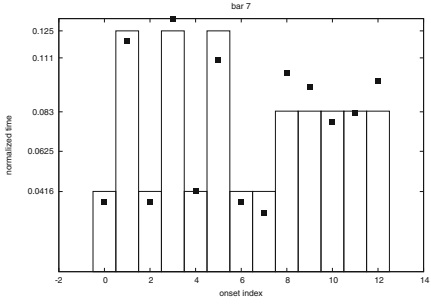












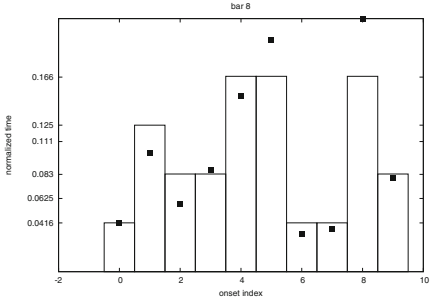
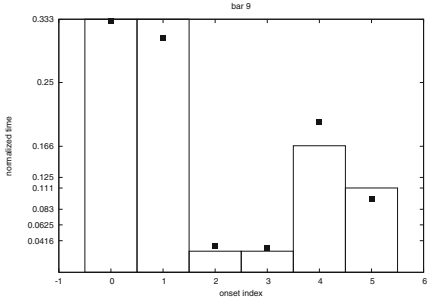
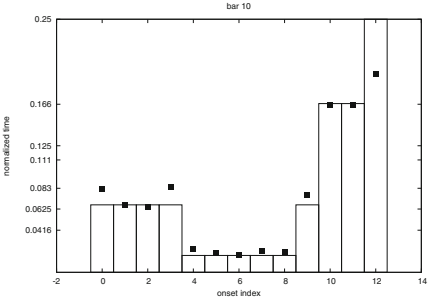
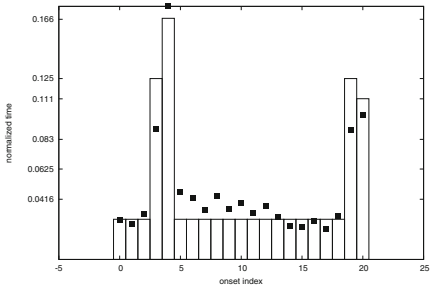
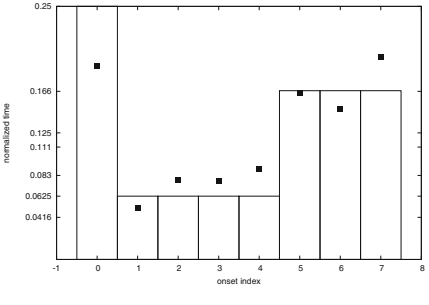


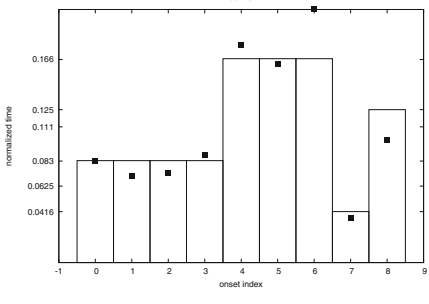
Table A.6 Quantization of Gould's 1955 recording of the Aria. Bars 9–16

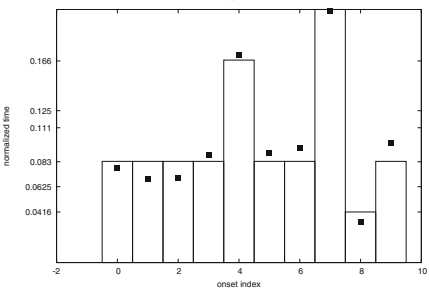


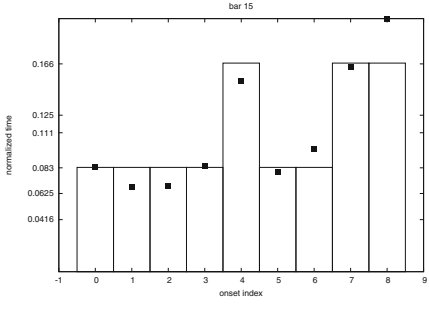












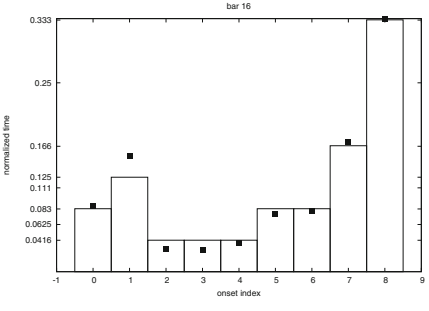


Table A.7 Quantization of Gould's 1955 recording of the Aria. Bars 17–24

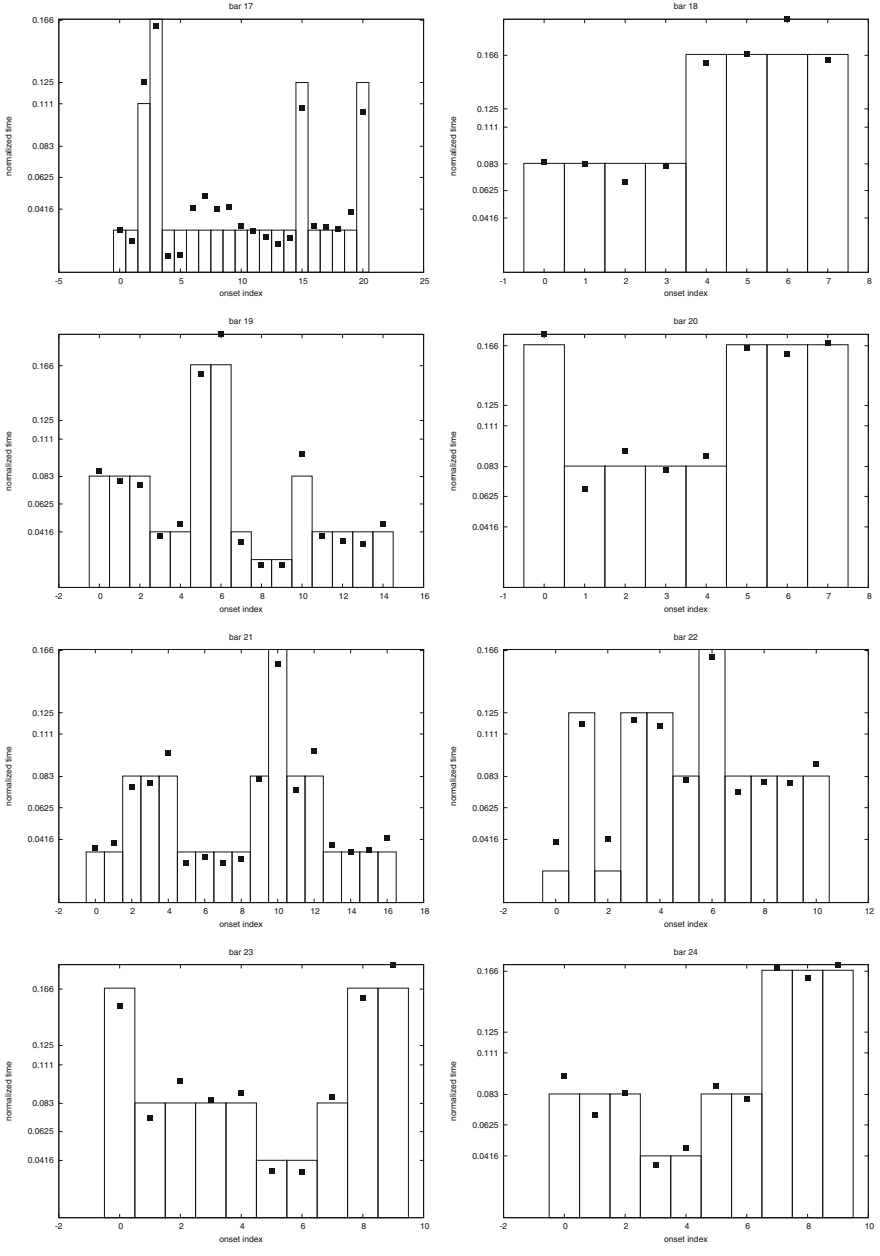
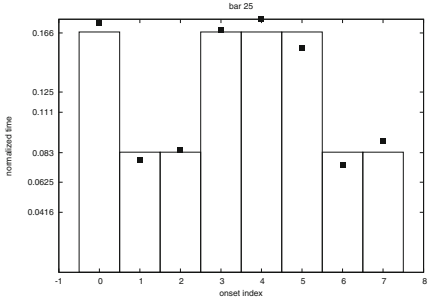
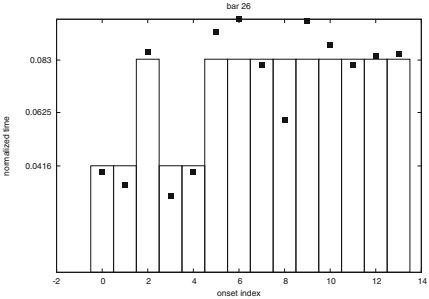
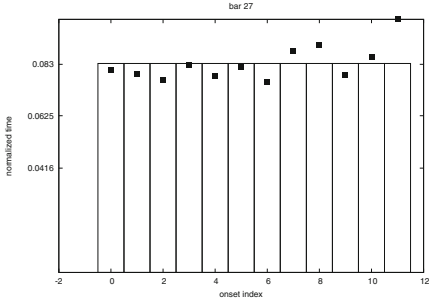
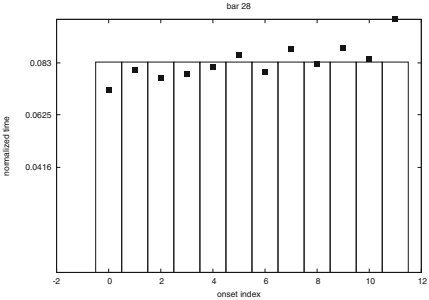


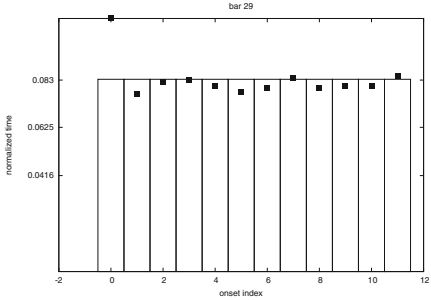
Table A.8 Quantization of Gould's 1955 recording of the Aria. Bars 25–32

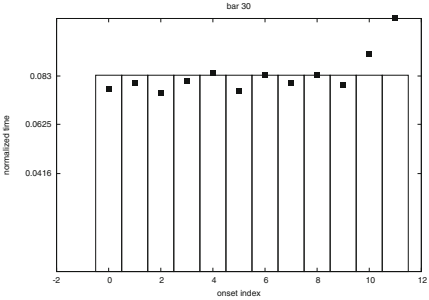


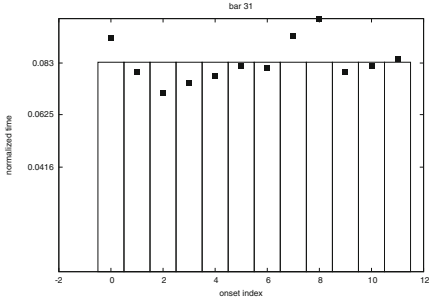


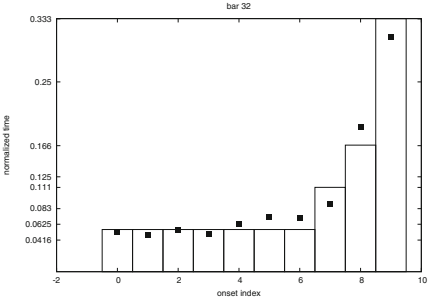












Index

A
Adjacent interval spectrum, 124, 142
African polyrhythm, 99
Agogics, 127
Aksak rhythm, 33, 38, 53, 75
Al-Urmawī, Safī al-Dīn, 115
Arab rhythms, 115
Articulation, 118
Assayag, Gérard, 138
Asymmetry, 39
Audio feature extraction, 124
Augmentation, 100

B
Bach, Carl Philip Emanuel Bach, 127
Bach, J. S., 171
  Goldberg Variations, 114, 159
Barlow, Clarence
  harmonicity function, 106, 158
  indigestibility function, 44, 104, 153
Bartók, Béla
  Musik for String Instruments, 46
Bayesian statistics, 136, 140
Bayes' theorem, 133
Beat, 9, 114, 117, 121
Beat and tempo tracking, 139
Beethoven, Ludwig van, 27
  Coriolan overture, 47
  Piano sonata no. 13, c-minor, 47
Berg, Alban
  violin concerto, 48
  Wozzeck, 46
Bernstein, Leonard, 135
Bioinformatics, 57
Braclets, 51
Breathing, 12
Brossier, Paul M., 124, 126
Brower, Candace, 12
Bruckner, Anton, 48
Bullock, Jamie, 124
Burrows–Wheeler Transform, 57, 69
  inverse, 58, 60
  matrix, 58

C
Categories of rhythmic patters, 52
Causal continuity, 7, 36, 127
Cause and effect, 7
Celibidache, Sergiu, 11
Cemgil, Ali Taylan, 136, 140
Chomsky, Noam, 135
Chopin, Fréderic, 128
Christoffel rhythm, 65
Christoffel word, 65
  palindrome, 66
Chunking (cognition), 18, 39
Chunking (software), 15, 33, 61
  sentence algorithm, 49, 53
Computer Music, 99
Conjugates, 58, 67
Context-free grammar, 135
Correlation, 9
Cycle, 115, 119
  in organic form, 11

D
Debussy, Claude, 29
Deutsch, Diana, 129, 130
Diminution, 100
Dixon, 139
Dupré, Marcel, 25

Duration classes, 125, 148
Dynamic function, 126, 127

E
Efrati, Richard R., 127, 129
Ellipsis, 100
Entrainment, *115*, 117, 120
Essen corpus of folksongs, 133
Euclidean distance, 138, 153, 160, 164
Euclidean rhythms, 77
Euclidean strings, 67
Euler's totient function, 86
Ewe rhythm, 23, 73, 115
Expressive timing, 126, 148, 160, 171, 172

F
Farey sequence, *84*
  filtered, 88, 102, 149, 151, 156
Final ritardando, 171
Fineberg, Joshua, 138
Force-Sensitive Resistor (FSR), 123

G
Gagaku, 125
Gamelan, 125
Gaussian window, 153
Gestalt rules, 128, 129, 142
Giger, Peter, 15, 18
Gould, Glenn, 125, 149, 152, 159, 164
  singing, 12
Greek verse rhythms, 25, 41, 72, 100
Grosche, Peter, 124
Growth and decline, 10
Gustafson, Kjell, 124

H
Hancock, Herbie, 155
Harmony, 130
100 milliseconds threshold, 121, 149, 171
Husserl, Edmund, 7

I
Image schema, 12
Impulse and release, 7
Integer partitions, 34
  into k distinct parts, 35, 49
  of the N-cycle, 109
  permutations, 34
  standard deviation, 35
  with arithmetic progression, 36
IRCAM, 124
  KANT, 137

K
Kilian, 135, 140
Köhler, Wolfgang, 128
K-smooth numbers, 92, 156

L
Langer, Susanne K., 10, 11
Large, Ed, 131
Latency, 123
Latin-American rhythms, 24, 73
Ligeti, György, 95
London, Justin, 38, 121, 142
  Many Meters Hypothesis, 119

M
Macedonian folk music, 75
Mahler, Gustav, 48
Manual tapping, 122
Markevitch, Igor, 25
Measure of metric complexity, 38
  algorithm, 44
Messiaen, Olivier, 12, 18, 25, 33, 86
  bird songs, 41
  Chronochromie, 41
Meter, 16, 35, 36, 114
  definition, 9
  hierarchy, 114
  non-isochronous, 38
Metric shift, 23, 61
  examples, 20
Metronome, 114
Miller, George Armitage, 39, 53
MPEG-7, 124
Müller, Meinard, 124
Musical form, 10
Musical triangle, 8, 38
  image schema of, 12
Mussorgsky, Modest, 28

N
Necklace, 67, 119
Notation, 15
Notes inégales, 168, 171
Novelty curve, 124

O
Ockeghem, Johannes, 95
Onset detection, 122, 124, 148
Organic form, 10, 38
Organism, 10
Ornaments, 121, 142, 149, 159, 160, 171

P
Pattern matching, 140
Phenomenology, 7
Physical durations, 118
Polak, Rainer, 36
Polyrhythm, 9, 29, 85, 118
  and Farey Sequences, 93
Psychoacoustics, 126
Psychological present, 122
Pulsation, 124
Pulse, 124, 131

Q
Quantization, 134
Quantz, Johann Joachim, 127

R
Raphael, Christopher, 136, 140
Rasch, Rudolf A., 130
Rhythm
  definition, 8
  transformations, 100
Rhythmic oddity, 75
Rhythmoglyphs, 18

S
Sapp, Craig Stuart, 124
Schoenberg, Arnold, 10
Shorthand notation, 15, 53
  table of symbols, 20
Stern–Brocot tree, *88*
Stockhausen, Karlheinz, 115
Stravinsky, Igor, 33
  Histoire du Soldat, 39
  Symphonies of Wind Instruments., 40
  The Rite of Spring, 97
Subdivisions, 20
  table of, 20
Suffix array, 59
Syncopation, 9, 35

T
Tchaikovsky, Peter Illich, 48
Temperley, 132
Template, 39, 50
Tempo, 118
Temporal expectancies, 118
Tempo rubato, 127, 128
Thakar, Markand, 7, 126, 127
Timbre, 126
Tosi, Pier Francesco, 127
Transcription, 53
Transients, 126
Triangular rhythm, 48
Turkish classical music, 75

V
von Ehrenfels, Christian, 129

W
Wagner, Richard, 48
Wangelin–Buschmann, Elisabeth, 128
Weakly Connected Neural Networks (WCNN), 131
Word
  as a rhythmic pattern, 58, 61
  Christoffel, 65
  definition, 2
  Lyndon, 67
  rotations, 58